(1)

AFIT/GST/ENS/93M-01

AD-A262 485

AN ANALYSIS OF THE

GENERALIZED LAMBDA DISTRIBUTION

THESIS

Robert J. Bergevin
First Lieutenant, USAF

AFIT/GST/ENS/93M-01

93-06836

Approved for public release; distribution unlimited

93 4 02 006

20001026165

1

AN ANALYSIS OF THE

GENERALIZED LAMBDA DISTRIBUTION

THESIS

Robert J. Bergevin
First Lieutenant, USAF

AFIT/GST/ENS/93M-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

THESIS APPROVAL

STUDENT: 1LT Robert J. Bergevin

CLASS: GST-93M-01

THESIS TITLE: An Analysis of the Generalized Lambda Distribution

DEFENSE DATE: 3 March 1993

COMMITTEE:	NAME/DEPARTMENT	SIGNATURE
Advisor	Dr. Edward Mykytka/ENS	Edward F. Mykytka
Reader	LTC Paul Auclair/ENS	Paul F. Auclair

Acknowledgements

There are several individuals who have been extremely helpful to me during the course of this research. My thesis committee members, Dr. Edward Mykytka and LTC Paul Auclair, were especially valuable. They provided me with numerous suggestions and improvements, as well as the encouragement to continue when things were not going well. I would have been lost without their help

I would also like to thank Art Earl for his help with the LaTeX document preparation system. His expertise with both LaTeX and the Sun Workstation operating system made the writing of this thesis much easier.

Finally, I would like to thank Debbie for her patience and understanding. I couldn't have accomplished this without her help and support.

Robert J. Bergevin

Table of Contents

List of Figures

List of Tables

Abstract

The Generalized Lambda Distribution (GLD) is a four parameter function that is capable of mimicking the behavior of a wide range of probability density functions (pdfs). Unfortunately, the GLD presently cannot model every possible type of pdf. Since the reasons for this limitation are unknown, this thesis examines several potential problems in an attempt to expand the range of distributions the GLD can mimic.

We first present a discussion of the behavior of the algorithm (known as Powell's Algorithm) that is used to search for the appropriate GLD parameter values. In particular, we examine the effect of using this unconstrained search procedure to find the parameters subject to a constraint that ensures that the resulting pdf is valid. We also develop a reparameterization of the GLD that creates an unconstrained search region. This does not expand the range of distributions the GLD can mimic.

We then use an extensive numerical investigation to examine the set of distributions that can be obtained from combinations of the GLD parameters. This examination allows us to expand the range of pdfs that can be modeled using the GLD. We also inspect some pdfs that cannot be modeled using the GLD, as well as present an alternative to the method of moments for determining the parameter values, using the recently-developed concept of L-moments.

AN ANALYSIS OF THE GENERALIZED LAMBDA DISTRIBUTION

I. Introduction

Today's military is faced with a two-fold problem. Because of the end of the Cold War and its resulting budgetary reductions, it must reduce its manpower in all career fields. In addition, since there is no longer a clearly-defined "enemy," it must also be prepared to deal with a myriad of potential conflicts and crises. For example, the military support for Operation Restore Hope, the 1992-93 relief efforts in Somalia, had to be thoroughly planned and organized. Military planners were faced with the monumental task of orchestrating the movement of over 28,000 Marine and Army troops, as well as providing for their logistical support in a country that had no established government, infrastructure or means of feeding its own people.

Because of budget cuts, the planning involved in dealing with such situations must be accomplished by a smaller staff, in the same amount of time as before. Planners are forced to do more with less and must therefore work more efficiently. Any cost– or time–cutting measures that are available must be implemented.

One such measure is the use of computer-aided simulations. As a result of the rapid improvement in the capabilities of small computer systems, computer simulations that model changing, complex, or unique situations have become more practical and easier to implement. Planning for contingencies such as Operation Restore Hope can be done in a fraction of the time, and the software used to analyze one scenario can be saved to model any other similar world situation that may arise in the future.

Simulation is useful in many military situations. It can be used to assess potential casualty rates and mission timetables for large–scale military operations, such as Desert Storm, the 1992

Gulf War. It can also be used to examine situations that are difficult to model in the real world, such as the expected damage resulting from an ICBM strike.

The problem with simulation, however, is that many applications require one to develop probabilistic models of various input variables, usually based on real-world data. For these simulations, information such as the lifetime of a certain part, the number of troops needed to secure an airport, or the radiation pattern from a nuclear explosion is crucial in order to get results that "mean something." However, such information is rarely known with certainty. Instead, planners usually have a bank of accumulated data from previous tests and experiences which they can use to probabilistically describe the range of possible values for each of their variables of interest.

The probabilistic description of a continuous random variable is usually summarized by its probability density function (pdf). Many different pdfs, including those for such well-known distributions as the Normal, Exponential, and Weibull, are available to the modeler. Each pdf has its own distinct shape (or shapes) and covers a specific range of values. Since most of these pdfs are available in commercial simulation packages, the modeler must decide which of the various distributions that are available best "fits," or describes, his set of data.

An alternative to this decision is to use a distribution with a more generalized density function that is not limited to a small number of particular shapes, but can instead take on a wide variety of different ones. The use of such a generalized function frees the modeler from having to decide between two competing distributions and allows him the leeway to create an even better fit to the data.

The Generalized Lambda Distribution (GLD) has just such a density function. It was originally created by Ramberg and Schmeiser [12] for the purpose of efficiently modeling and generating random variables for use in simulation studies. The GLD is comprised of four parameters, λ_1, λ_2, λ_3, and λ_4, that act together to allow adjustment of the location, scale, and shape of the density function so that it is able to model different random variables. In order to fit the GLD to an actual set of data, we need to determine appropriate values for the four parameters. This is usually accomplished by first computing the first four sample moments (the mean, variance, skewness, and kurtosis) of the data. We then use a computerized search routine to find the appropriate combination of parameter values so that the resulting form of the GLD will have those same four moments.

Because of its ability to "match moments," we can use the GLD to mimic the behavior of most of the commonly-used pdfs simply by setting the four parameters to appropriate values. Figure 1 shows the ranges of these pdfs in terms of their measures of skewness and kurtosis. Some

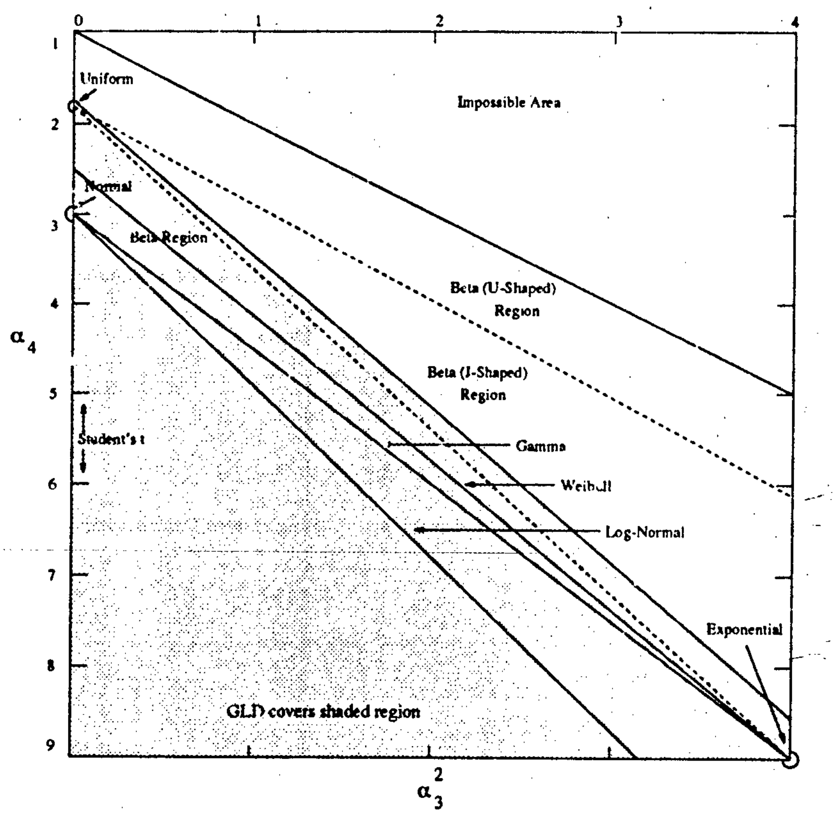

Figure 1. Characterization of Various Distributions by Their Skewness and Kurtosis

distributions, such as the Uniform, Normal and Exponential, are represented by single points; others, such as the Student's t, Log-Normal and Gamma, are represented by curves; the three types of beta distributions are represented by regions of values. The top right-hand region, denoted the "Impossible Area," contains skewness-kurtosis combinations that will never be exhibited by any pdf. The GLD, at present, can "mimic" those pdfs found within the shaded region of Figure 1. Obviously, all but some of the U- and J-shaped beta distributions can currently be modeled using the GLD.

In order to make the GLD a more effective tool, we wish to expand the shaded region of Figure 1 beyond its current limits—ideally, all the way to the Impossible Area. Unfortunately, the reasons for the GLD's current limitation are unknown, and one of the objectives of this thesis is to explore these limitations and develop means to overcome them.

There are three particular questions we can address in pursuit of this objective. One concerns the behavior of the computerized search routine (known as Powell's Algorithm) used in the moment-matching process. This algorithm was originally designed for searches over unrestricted regions. However, in the case of the GLD, two of the parameters, λ_3 and λ_4, are required to have the same sign, thereby limiting the range of possible values. We are uncertain of the behavior of our particular implementation of Powell's Algorithm as it encounters this restricted area. In order to get a better understanding of this problem, we will perform an examination of the algorithm's behavior as it approaches and attempts to cross the constraining boundaries. We will also examine some techniques for eliminating the problem altogether. These efforts should help us determine whether the observed limitation on the GLD occurs as the result of inadequacies in the implementation of this search algorithm.

Second, previous research has shown that a pattern exists in the values of λ_3 and λ_4 that produce certain combinations of skewness and kurtosis. This pattern suggests the hypothesis that the minimum possible value of kurtosis that the GLD can have—given a specified value of skewness—occurs when $\lambda_3 = 0$. This implies that the GLD's limitation might be a result of restrictions on the possible values of its parameters. As a means of assessing the validity of this hypothesis, we will examine the skewness and kurtosis values that result from different combinations of λ_3 and λ_4. We hope that the insights garnered from this exercise will also enable us to determine if the limitations on the GLD can be overcome.

Third, we will examine some of the distributions that can not be modeled by the GLD. Although these cases may be important in mathematical circles, it is conjectured that they may not have significant practical applications. Further examination will help determine whether this is true or not.

As discussed previously, the GLD parameter values are traditionally estimated from the first four moments of the data. However, the higher order sample moments (skewness and kurtosis) are highly variable and are, therefore, somewhat untrustworthy. Recent research has suggested that a different approach. The use of L-moments could be used to replace these moments with potentially more–stable measures. Thus, a second objective of this thesis is to develop these measures for the GLD and revise the computer search routines to utilize this new method as an alternative to the method of moments.

Figure 1 shows that the GLD can already be used as an effective tool for simulation analyses. Hopefully this research will provide a deeper insight into the properties and limitations of the GLD. By conclusively determining the limits of its range, as well as the types of distributions that can not be duplicated using it, we will attempt to make the GLD an even more powerful tool and further enhance its value for future simulation studies.

II. Background

2.1 GLD Development

Although most of the commonly-used continuous probability distributions are defined in terms of their density functions, f(x), or cumulative distribution functions, F(x), it is equally valid to define a distribution by its percentile function, if that percentile function exists. The percentile function is simply the inverse of the distribution function, i.e., $R(p) = F^{-1}(p)$, or equivalently, $p = F(x)$. The percentile function, R(p), is used similarly to a distribution's cumulative density function (cdf), in that it determines the value, $x = R(p)$, such that the probability that a random variable having this distribution takes a value less than x is p.

The ability to express a random variable in terms of its percentile function is quite useful in Monte Carlo simulation studies. In particular, it is well known that if R is the percentile function of a continuous probability distribution and if the random variable U is uniformly distributed on the (0,1) interval, then the transformation $X = R(U)$ yields a continuous random variable with percentile function R. Thus, since sources of uniform (0,1) pseudo-random variates are commonly available, this transformation yields a simple method for generating pseudo-random variates from distributions whose percentile functions are known and are computationally tractable.

Tukey [15] created a function, which he called the lambda function, in this manner. Tukey's function, which is valid for all non-zero λ, can be written as

$$R(p) = \frac{p^\lambda - (1-p)^\lambda}{\lambda}, \quad 0 \leq p \leq 1. \tag{1}$$

Filliben [2] used this function to approximate symmetric distributions with a wide range of tailweights and noted that when $\lambda = 0.14$, the lambda function resulted in a "good" approximation to the standard normal distribution. He further noted that the logistic distribution results as

$\lambda \to 0$ while, for $\lambda = -1$, the resulting function is approximately Cauchy. Filliben also presented a complete discussion of the density functions that result for various values of λ.

Ramberg, *et al* [13] generalized Equation 1 to a four-parameter distribution that could be used to approximate a number of well-known symmetric and asymmetric distributions—noting, in comparison, that a close approximation to the standard normal results when $\lambda_1 = 0$, $\lambda_2 = 0.1975$, and $\lambda_3 = \lambda_4 = 0.1350$ [13:203]. Their distribution is defined by the percentile function

$$R(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2}, \quad 0 \leq p \leq 1. \tag{2}$$

The distribution defined by Equation 2 is referred to as the Generalized Lambda Distribution (GLD). The GLD has also been referred to as the Ramberg-Schmeiser-Tukey (RST) distribution in the literature (see, for example, Mykytka [9]). The parameters λ_1 and λ_2 are location and scale parameters, respectively, while λ_3 and λ_4 are shape parameters that jointly determine the skewness and kurtosis of the GLD. When $\lambda_3 = \lambda_4$, the resulting density is symmetric.

Using the fact that $x = F^{-1}(p) = R(p)$, we can find the density function corresponding to Equation 2 by noting

$$f(x) = \frac{dF(x)}{dx} = \frac{dp}{dR(p)} = \left(\frac{dR(p)}{dp}\right)^{-1};$$

which yields

$$f(x) = \left(\frac{dR(p)}{dp}\right)^{-1} = \frac{\lambda_2}{\lambda_3 p^{\lambda_3 - 1} + \lambda_4 (1-p)^{\lambda_4 - 1}} \quad 0 \leq p \leq 1. \tag{3}$$

It should be noted that although λ_1 does not appear explicitly in this expression, f(x) is indeed a function of λ_1 since it is defined in terms of R(p), which does depend on λ_1.

The cumulative distribution function of the GLD does not, in general, exist in a simple closed form, but this should not be a cause of concern since it is also true of the normal distribution, whose percentiles are more difficult to compute. For the GLD, it is simple to obtain plots of the distribution function by plotting p on the y-axis versus R(p) on the x-axis. Similarly, a plot of the density function is obtained by plotting f[R(p)] on the y-axis against R(p) on the x-axis, for p ranging from zero to one. FORTRAN programs that compute R(p) and f[R(p)] for specified lambda values are given in Mykytka [9:82-84].

2.2 Calculation of GLD Parameters

2.2.1 Statistics Background. The four GLD parameters are linked to the distribution's first four central moments: the mean (μ), variance (σ^2), skewness (α_3), and kurtosis (α_4). For readers unfamiliar with these concepts, this subsection will present a brief overview. A more thorough explanation of these concepts can be found in most statistics textbooks. The information given below was taken from Mendenhall, Wackerly and Scheaffer [8].

The first moment about the origin describes the center of a pdf and is commonly referred to as the mean (μ). The variance (σ^2) is the second moment about the mean of a distribution and describes the "spread" if its pdf. Unfortunately, the mean and variance do not uniquely define a pdf. Many different distributions can possess the same mean and variance. Therefore, we must utilize additional measures (such as skewness and kurtosis) to distinguish between different pdfs.

The reader might remember that for a particular distribution, the mean is defined as the expected value of a random variable following that distribution, $\mu = E(X)$. This is the average value we would observe if we were to continually take samples from the distribution. For an empirical data set with n observations, we can estimate μ with the sample mean, $\overline{X}$, which is defined as follows

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

$\bar{X}$ yields the average value of the data set.

The variance gives a measure of how wide the distribution (or sample data set) is. It is defined as the expected value of the square of the difference between a sample value and the mean of the distribution, $\sigma^2 = E[(X - \mu)^2]$. For sample data, this can be estimated using the sample variance $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

The higher order moments are defined similarly. The standardized third moment about the mean, the skewness (α_3), gives a measure of how symmetric the distribution is. It is defined as the scaled expected value of the cube of the difference between a sample value and the mean of the distribution, $\alpha_3 = \frac{E[(X-\mu)^3]}{\sigma^3}$, where the σ^3 term in the denominator is a scaling factor used to make α_3 a dimensionless measure. For sample data sets,

$$\hat{\alpha_3} = \frac{1}{n}\frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{\hat{\sigma}^3}.$$

If a distribution is symmetric about its mean (like the Normal distribution, for example), it will have a skewness of zero.

The standardized fourth moment about the mean, the kurtosis (α_4), is a measure of the "tailweight" of the distribution. It can be roughly thought of as the number of values that lie in the tails of a distribution. It is defined as the scaled expected value of the difference between a sample value and the mean, taken to the fourth power, $\alpha_4 = \frac{E[(X-\mu)^4]}{\sigma^4}$. We again use a scaling factor (σ^4) to make α_4 a dimensionless measure. For sample data sets, α_4 can be estimated using:

$$\hat{\alpha_4} = \frac{1}{n}\frac{\sum_{i=1}^{n}(X_i - \overline{X})^4}{\hat{\sigma}^4}.$$

A lower value of α_4 signifies that the distribution will have "thinner" tails than a distribution that possesses a larger value of α_4.

As mentioned previously, many different distributions can share the same mean and variance. Fortunately, a distribution's measures of skewness and kurtosis are fairly unique, and therefore can be used to distinguish it from other pdfs, as we saw in Figure 1. Although we can only uniquely define a distribution using an infinite number of its moments, we will find that, in practice, using just these first four moments will be sufficient for most purposes.

These four moments will be used to determine values for the four parameters of the GLD. By setting the four λ parameters to appropriate values, we will be able to duplicate many combinations of mean, variance, skewness, and kurtosis.

2.2.2 Parameter Definitions. Ramberg and Schmeiser [12] showed that for $\lambda_1 = 0$, the k^{th} moment of the GLD, when it exists, is given by

$$E(X^k) = \lambda_2^{-k} \sum_{i=0}^{k} \binom{k}{i} (-1)^i \beta(\lambda_3(k-i)+1, \lambda_4 \cdot i + 1) \tag{4}$$

where the beta function is defined (as in [1:332]) by

$$\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Since the beta function is undefined whenever either of its arguments are negative, we must assure both:

$$\lambda_3(k-i) + 1 \geq 0$$

and

$$\lambda_4 i + 1 \geq 0,$$

for all $i \leq k$. Therefore, the k^{th} moment does not exist whenever

$$min(\lambda_3, \lambda_4) < -\frac{1}{k}. \qquad (5)$$

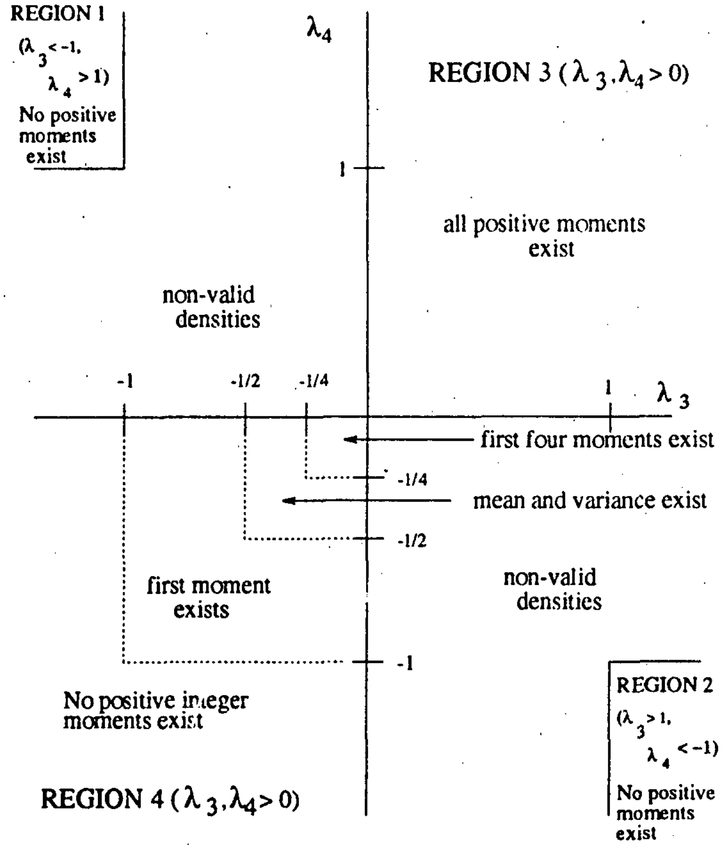

Figure 2. Regions of GLD in λ_3–λ_4 Space

The mathematical definition of the GLD that has been given does not ensure that it always defines a legitimate probability distribution. A valid pdf must exhibit the following behavior:

$$f(x) \geq 0, \qquad \forall x$$

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Schmeiser [14] determined that there are four regions for which the GLD exhibits such behavior. These regions are depicted in λ_3–λ_4 space in Figure 2, and have been arbitrarily numbered 1, 2, 3, and 4. It can be shown that no positive moments exist in Regions 1 or 2 (see, for example, [6:183-193]). Likewise, from equation 5, we can see that the first four moments exist in Region 4 only when both λ_3 and λ_4 are greater than $-\frac{1}{4}$. All positive moments exist in Region 3.

Using Equation 4, Ramberg and Schmeiser developed the following expressions for the mean, variance, skewness, and kurtosis of the GLD:

$$\mu = \mu(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{A}{\lambda_2} \tag{6}$$

$$\sigma^2 = \sigma^2(\lambda_2, \lambda_3, \lambda_4) = \frac{B - A^2}{{\lambda_2}^2} \tag{7}$$

$$\alpha_3 = \alpha^3(\lambda_3, \lambda_4) = sign(\lambda_2) \cdot \frac{C - 3AB + 2A^3}{(B - A^2)^{\frac{3}{2}}} \tag{8}$$

$$\alpha_4 = \alpha^4(\lambda_3, \lambda_4) = \frac{D - 4AC + 6A^2B - 3A^4}{(B - A^2)^2} \tag{9}$$

where A, B, C, and D are the following functions of λ_3 and λ_4:

$$A = \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4} \tag{10}$$

$$B = \frac{1}{1+2\lambda_3} - 2\beta(1+\lambda_3, 1+\lambda_4) + \frac{1}{1+2\lambda_4} \tag{11}$$

$$C = \frac{1}{1+3\lambda_3} - 3\beta(1+2\lambda_3, 1+\lambda_4) + 3\beta(1+\lambda_3, 1+2\lambda_4) - \frac{1}{1+3\lambda_4} \tag{12}$$

$$D = \frac{1}{1+4\lambda_3} - 4\beta(1+3\lambda_3, 1+\lambda_4) + 6\beta(1+2\lambda_3, 1+2\lambda_4)$$
$$-4\beta(1+\lambda_3, 1+3\lambda_4) + \frac{1}{1+4\lambda_4} \quad (13)$$

and $sign(\lambda_2)$ will be either 1 or -1, depending on whether the λ_2 parameter is positive or negative.

As the notation indicates, the skewness and kurtosis are functions of λ_3 and λ_4 alone. The skewness, α_3, is a function of only λ_3 and λ_4 since, in Regions 3 and 4, the sign of λ_2 is always the same as the sign of both λ_3 and λ_4 [9:20]. The variance, however, also depends on the shape parameter λ_2 and the mean depends on all four parameters. The shaded region of Figure 1 shows the different combinations of skewness and kurtosis that can be obtained from the GLD.

2.2.3 Techniques for Determining Parameter Values. The technique known as the method of moments is the usual means of selecting the values of the GLD parameters. By choosing the four parameter values appropriately, a wide range of distributions can be duplicated, as indicated by the shaded portion of Figure 1. We choose the parameter values as follows:

1. Since the equations for the skewness and kurtosis, Equations 8 and 9, are functions of only λ_3 and λ_4, we first determine the values of λ_3 and λ_4 that "match" the desired combination of α_3 and α_4.
2. Since we now have values for λ_3 and λ_4, we use Equation 7 to find a value for λ_2 so that the GLD has the desired variance.
3. Equation 6 is then used to find a value for λ_1 so that we achieve the desired mean as well.

Table 1. Parameter Determination Using the Method of Moments

If the values for μ, σ^2, α_3, and α_4 are unknown, they must be estimated from the sample data. The procedure outlined above can then used to determine the four lambda parameters, replacing μ, σ^2, α_3, and α_4 with the sample statistics $\overline{X}$, $\hat{\sigma^2}$, $\hat{\alpha_3}$, and $\hat{\alpha_4}$.

It should be noted that the third and fourth sample moments, $\hat{\alpha_3}$, and $\hat{\alpha_4}$, are quite sensitive to extreme observations (values of X_i located more than two standard deviations from $\overline{X}$) and the

variability of these sample moments can be quite large. This can, in turn, result in a poor, or even incorrect, GLD fit. As a way of reducing this variability, Hosking [3] presents an alternative to the use of sample moments, that of L-moments. His L-moments are computed as linear combinations of order statistics. When properly defined, these L-moments can be used in place of the traditional sample moments. He has shown that L-moments produce a more powerful goodness-of-fit test of Normality than do traditional moments [4]. We will consider the use of L–moments in selecting the parameters of the GLD in Chapter VII.

Ramberg *et al* [13:210-214] provide a table of the four GLD parameter values for various combinations of skewness and kurtosis. These tables only cover distributions with zero mean and unit variance, but the transformations

$$\begin{aligned} \lambda_1(\mu,\sigma^2) &= \lambda_1(0,1)\cdot\sigma+\mu \\ \lambda_2(\mu,\sigma^2) &= \frac{\lambda_2(0,1)}{\sigma} \end{aligned}$$

can easily be computed for cases when $\mu \neq 0$ and/or $\sigma^2 \neq 1$.

The calculations for finding the values of λ_3 and λ_4 via the method of moments, given specified values of α_3 and α_4, are complicated. Several different techniques have been implemented. Mykytka and Ramberg [10] and Mykytka [9] use non–linear programming methods to find the minimum possible sum-of-squared errors between the calculated and desired values of α_3 and α_4:

$$Min \quad f(\lambda_3,\lambda_4) = (\alpha_3(\lambda_3,\lambda_4)-\alpha_3)^2+(\alpha_4(\lambda_3,\lambda_4)-\alpha_4)^2 \tag{14}$$

$$subject\ to \qquad \lambda_3\cdot\lambda_4 \geq 0. \tag{15}$$

Equation 15 insures that λ_3 and λ_4 lie in either Region 3 or Region 4. The minimization expressed in Equation 14 is performed using Powell's Algorithm for non–linear function minimization.

The tables of Ramberg *et al* were originally developed by specifying desired values for α_3 and α_4, then finding the appropriate values of λ_3 and λ_4 using a FORTRAN program to solve the minimization problem described in Equation 14. These tables contain values of the four lambda parameters for values of α_3 between zero and two (in increments of 0.1) and values of α_4 in increments of 0.2. For a given value of α_3, the values for α_4 tabled are the smallest values for which the optimal value of the objective function, Equation 14, was approximately zero. Thus, for a given value of α_3, the table does not necessarily show the minimum possible value of α_4 that is theoretically possible, but only that for which an objective function value near zero was obtained.

Mykytka and Ramberg [10] provide a user-friendly FORTRAN program that will calculate the four GLD parameter values for combinations of skewness and kurtosis not given in the tables of [13]. In addition, Hsu [5] has created a C++ program which calculates the four GLD parameters for a given data set and then allows the user to visually "improve" the resulting fit to the data by altering the value of one or more of the parameters.

III. Thesis Objectives

In order for the GLD to be an effective probability distribution, useful for modeling a wide variety of random variables, we desire that it to be able to "mimic" any possible combination of skewness and kurtosis. At present, it cannot do so. Mykytka [9] showed a tentative boundary (reproduced here in Figure 1) for the range of α_3–α_4 combinations that can be produced using the GLD. This boundary is based on his tabulated results, which are also found in Ramberg *et al* [13]. However, as mentioned previously, these tabulated results only show combinations of α_3 and α_4 for which the numerical search procedure produced a solution with a near-zero objective function. Therefore, we are not assured that additional combinations do not exist above this boundary.

We wish to take steps to either confirm Mykytka's boundary or to expand the GLD's coverage region. However, due to the complexity of the procedures used to find the λ_3 and λ_4 values that correspond to a specified skewness–kurtosis combination, there are several potential problems that could be limiting the range of distributions the GLD can mimic. Each of these problems will be discussed in detail in the following chapters, along with possible methods for their elimination. We may discover that none of these problems affect the GLD's coverage region. If that is indeed the case, the efforts outlined in this thesis will not be in vain. By thoroughly examining these concerns, we will firmly establish the limitations of the GLD, which will be of value to future researchers.

In Chapter IV, we present an in-depth analysis of the particular implementation of Powell's Algorithm used by Mykytka and Ramberg [10]. The algorithm was originally designed to search over an unconstrained region, so we will concentrate on its behavior as it approaches the constraints defined by Equation 15. We then examine the effects of the penalty function that is currently used to enforce these boundaries. Finally, we describe possible remedies to the problems involved with using Powell's Algorithm. We analyze one possible solution in particular: a reparameterization of the GLD parameters in order to create an unconstrained search region.

In Chapter V, we present an analysis of some possible limitations of the GLD. Instead of determining the values of λ_3 and λ_4 that correspond to a particular α_3–α_4 combination, we instead look at the values of α_3 and α_4 that result from different combinations of λ_3 and λ_4. By examining a wide range of λ_3–λ_4 combinations, we may be able to gain some insight about and expand the limits on the GLD's range.

Chapter VI examines some of the distributions that are not presently covered by the GLD. By examining these pdfs, we will get a better grasp for the types of distributions we can not yet model.

As mentioned previously, the variability of the higher-order sample moments can be quite large, especially in small data sets. Because of this potential problem, Chapter VII describes L-moments, an alternative to the use of the traditional sample moments. It provides a summary of the GLD's four L-moment equations, as well as some suggestions for measuring L-moments from a empirical data set.

Finally, Chapter VIII summarizes the results of this research and suggests some possible paths for follow-on efforts.

IV. Analysis of Powell's Algorithm

4.1 Background

The search procedure used by Mykytka and Ramberg [10] to determine the optimal values for λ_3 and λ_4 makes use of Powell's Algorithm. This algorithm was originally designed for unconstrained non-linear function minimization. However, the GLD is constrained by the requirement that λ_3 and λ_4 have the same sign:

$$\lambda_3 \cdot \lambda_4 \geq 0.$$

As previously mentioned, our objective is to minimize the sum-of-squared errors between the desired and calculated values of α_3 and α_4 subject to this constraint (see Equations 14 and 15). In the current implementation, "unacceptable" values—those combinations that do not satisfy Equation 15—are eliminated from consideration by replacing their sum-of-squared errors with a large penalty. We enforce this penalty by defining the objective function as

$$min\ Z = f(\lambda_3, \lambda_4)$$

where

$$\begin{aligned} Z &= [(\alpha_3(\lambda_3, \lambda_4) - \alpha_3)^2 + (\alpha_4(\lambda_3, \lambda_4) - \alpha_4)^2]; && \lambda_3 \cdot \lambda_4 \geq 0 \\ Z &= 10; && \lambda_3 \cdot \lambda_4 < 0. \end{aligned}$$

Since an appropriate "match" to the desired α_3 and α_4 values is produced only when the optimal value of the objective is zero (practically interpreted as having an objective function value less than some arbitrarily small value[1]), we can avoid "unacceptable" combinations of λ_3 and λ_4 by utilizing this penalty.

[1] For example, the FORTRAN code presented in Mykytka and Ramberg [10] will warn the user of an unacceptable match when the final value calculated for Z exceeds 0.0002.

Unfortunately, we do not know how Powell's Algorithm acts when faced with this constraint, *i.e.*, does the inclusion of the penalty force the search into inappropriate directions? Our concern is that the penalty might possibly cause the search algorithm to exhibit undesired or detrimental behavior, which might limit the range of potential α_3-α_4 combinations for which we can find solutions.

As noted earlier, the only valid combinations of λ_3 and λ_4 lie in Regions 3 and 4—regions that are mutually exclusive[2] (see Figure 2). What happens if our initial estimates of λ_3 and λ_4 are in Region 3 when the actual result lies in Region 4? As Powell's Algorithm proceeds towards (0,0) in the λ_3–λ_4 space, it could be faced with the problem of facing a restricted region in each potential direction of travel. What happens in this case? Can we be certain that the algorithm moves in the proper direction?

Some preliminary answers already exist. When using the FORTRAN program given by Mykytka and Ramberg [10], the user must input an initial "guess" at the values for λ_3 and λ_4. Since λ_3 and λ_4 must have the same sign, their initial values must be either both positive or both negative. It can be shown that even if the initial guess is in the wrong region (for example, an initial guess in Region 4 could be provided when the optimal solution actually lies in Region 3) the algorithm will sometimes switch regions to find the optimal solution. On the other hand, due to the nature of the algorithm's search technique (which will be discussed in detail in Section 4.3), the possibility of jumping like this directly from one acceptable region directly to the other is slight. That is, although the algorithm sometimes does switch regions, it usually does not. If, in practice, we do not achieve a satisfactory Z–value when starting in one region, we need simply apply the algorithm a second time, changing our starting point to one in the opposite region and hopefully producing a solution with a better objective function value.

[2] The point $\lambda_3 = \lambda_4 = 0$, although common to both regions, does not yield a valid GLD pdf.

This behavior suggests that we could see some momentary attempts by the algorithm to find potential solutions in the "forbidden zone" (*i.e.* combinations of λ_3 and λ_4 that lie outside of Regions 3 and 4 in Figure 2) as the algorithm attempts to cross from one region to the other. What affect might this have on the search algorithm?

The more important cases for this concern are those where the optimal λ_3–λ_4 solutions lie very close to one of the "forbidden zone" boundaries. If we assume that our initial guess lies in the proper region, we can reasonably expect that the magnitude of Z will be decreasing as Powell's Algorithm approaches its optimal solution (minimum Z–value). As the algorithm nears this optimal point, its search might reach into the "forbidden zone." The algorithm is suddenly faced with a "large" increase in the Z–value (namely, $Z = 10$) in that particular direction. Does this bias the search in any way? Such a problem might cause the search to proceed in an inappropriate direction or prevent it from converging to a solution that lies near the constraining boundaries. As a result, the algorithm might fail to converge to a solution with a non–zero objective function value and the particular α_3–α_4 combination would be considered to be infeasible, even though a valid solution does indeed exist.

4.2 Methodology

To investigate this problem, we will examine the behavior of the algorithm, step-by-step, as it proceeds near—or attempts to cross—the limiting boundaries. This will be accomplished by altering the FORTRAN code of Mykytka and Ramberg [10] to print each of the intermediate points in a particular search. By examining the intermediate steps generated by the algorithm as it proceeds to its final solution, we can get a more complete understanding of its behavior and a fuller confidence in its results.

We will also inspect the behavior of the algorithm when we provide a starting value in the "opposite" region. We know there are some cases where the algorithm does cross from one region to the other in searching for its optimal value. By comparing the results of these searches to those when starting points in the proper region are used, we will get an insight into the behavior of the algorithm as it attempts to search for a solution in, or in the direction of, the "forbidden zone."

In Section 4.4, we will examine cases where the optimal λ_3 and λ_4 solutions lie close to the constraining boundaries. By examining the results of searches performed with and without the penalty function, we will be able to evaluate the effect of the penalty function on the algorithm.

We will use the tabulated values of Ramberg *et al* [13:210-214] as a source of test values for this effort. In particular, we will be examining values near (or past) the point where their particular implementation of Powell's Algorithm failed to converge. Although we will use the FORTRAN code presented in Mykytka and Ramberg [10] instead of the code used to generate the tables in [13], the two versions are nearly identical and extensive experience with the code suggests that we can be confident that they will yield identical results.

4.3 Unconstrained Behavior of the Algorithm

As a starting point, we examine the search techniques used in Powell's Algorithm. By studying how the algorithm usually performs searches in an unconstrained space, we will hopefully be able predict its behavior when it is presented with limiting boundaries. We will use Powell's original paper, [11], as well as the FORTRAN code of Kuester and Mize [7], as references for this analysis.

Powell uses a variation of the method of minimizing a function of several variables by changing only one parameter at a time. His method uses conjugate search directions at each iteration, which results in a "fast" rate of convergence.

Each iteration of the algorithm consists of linear searches down n independent directions, $\xi_1, \xi_2, \ldots, \xi_n$, where n is the number of variables for which we desire values. We start from the best approximation of the minimum, $\mathbf{p}_0$. Initially, the search directions are chosen to be in the direction of the coordinate axes and $\mathbf{p}_0$ is simply our initial "guess" of the point that yields the optimal objective function value. Each iteration defines a new search direction, ξ and, if a test is passed, this new direction replaces one of the original search directions. A description of an iteration of the algorithm is given in Table 2.

1. For $r = 1, 2, \ldots, n$ calculate ν_r so that $f(\mathbf{p}_{r-1} + \nu_r \xi_r)$ is a minimum and define $\mathbf{p}_r = \mathbf{p}_{r-1} + \nu_r \xi_r$.
2. Find the integer, m, $1 \leq m \leq n$, so that $[f(\mathbf{p}_{m-1}) - f\mathbf{p}_m)]$ is a maximum, and define $\Delta = f(\mathbf{p}_{m-1}) - f(\mathbf{p}_n)$.
3. Calculate $f_3 = f(2\mathbf{p}_n - \mathbf{p}_0)$, and define $f_1 = f(\mathbf{p}_0)$ and $f_2 = f(\mathbf{p}_m)$.
4. If either $f_3 \geq f_1$ and/or $(f_1 - 2f_2 + f_3) \cdot (f_1 - f_2 - \Delta)^2 \geq \frac{1}{2}\Delta(f_1 - f_3)^2$, use the old directions, $\xi_1, \xi_2, \ldots, \xi_n$ for the next iteration and use $\mathbf{p}_n$ as the next $\mathbf{p}_0$, otherwise,
5. Defining $\xi = (\mathbf{p}_n - \mathbf{p}_0)$, calculate ν so that $f(\mathbf{p}_n + \nu\xi)$ is a minimum. Use $\xi_1, \xi_2, \ldots, \xi_{m-1}, \xi_{m+1}, \xi_{m+2}, \xi_n, \xi$ as the directions and $\mathbf{p}_n + \nu\xi$ as the starting point for the next iteration.

Table 2. Iteration Procedure for Powell's Algorithm

The process outlined above is a modification to his original method, requiring a larger number of iterations, but in [11], Powell states that is a valuable, and in some instances, essential modification. The criterion for convergence is given in Table 3 (taken from Powell [11:158]).

1. Apply the normal procedure until an iteration causes the change in each variable to be less that one–tenth of the required accuracy, denote the relevant point as **a**.
2. Increase every variable by ten times the required accuracy.
3. Apply the normal procedure until an iteration causes the change in every variable to be less than one–tenth of the required accuracy again. Denote the resultant point as **b**.
4. Find the point at which the function is minimized over the line through **a** and **b**, denote this point as **c**.
5. Assume ultimate convergence if the components of $(\mathbf{a} - \mathbf{c})$ and $(\mathbf{b} - \mathbf{c})$ are all less than one–tenth the required accuracy in the corresponding variables, otherwise
6. Include the direction $(\mathbf{a} - \mathbf{c})$ in place of ξ_1, and restart the procedure from Step 1.

Table 3. Criterion for Convergence of Powell's Algorithm

In the FORTRAN code of Kuester and Mize [7] (which is the same code used by Mykytka [9] and Mykytka and Ramberg [10]), the linear searches are performed using quadratic approximation techniques (see Chapter 7 of [7]). The function is evaluated at three different points: $f(\mathbf{p})$, $f(\mathbf{p}+q\xi)$, and either $f(\mathbf{p}+2q\xi)$ or $f(\mathbf{p}-q\xi)$, depending on whether $f(\mathbf{p})$ is less than or greater than $f(\mathbf{p}+q\xi)$.

The term q represents the length of the step along the line, and $\mathbf{p}$ represents the current point in the search process. These three points are used to determine whether the function is at a local minimum. If it is not, the three points are used to generate the "turning value," which is calculated using a quadratic function of the points and their respective function values. The "turning value" determines in which direction the search will continue.

In our particular case, we should not expect the algorithm to venture into the "forbidden zone." If one of the algorithm's quadratic approximation searches attempts to enter this area, the large value for the penalty function should cause it to reverse course and head in the opposite direction—back into one of the valid regions. If we only searched in the directions of the coordinate axes, it would be almost impossible for the algorithm to begin a search in one region and conclude in the other. However, since the search directions will most likely change as we iterate through the process, it is possible that the algorithm will "switch" regions, although it appears this behavior will occur only in cases where Powell's Algorithm can move directly from one valid region to the other during the course of a single quadratic approximation search.

To test this assumption, we examined several cases where our starting point was in the wrong region, and the solution algorithm switched to the appropriate region for the optimal result. In all the cases that were analyzed, some common behaviors were observed. First of all, the algorithm always switched regions near the (0,0) point in the λ_3–λ_4 space. Secondly, none of the quadratic approximation searches terminated in the forbidden zone. In all cases, one particular quadratic approximation search had a starting point in the initial region and an ending point in the opposite region. The algorithm continued its search from this new point and eventually found the optimal solution.

We present the results of one particular case here as an example. For this analysis, we used the FORTRAN code of Mykytka and Ramberg [10] to find the lambda parameters for the following combination of the first four moments: $\mu = 0.0$, $\sigma^2 = 1.0$, $\alpha_3 = 0.0$, $\alpha_4 = 6.0$. This combination is included in the tables of Ramberg *et al* [13] with the following results: $\lambda_1 = 0$, $\lambda_2 = -0.1686$, $\lambda_3 = -0.0802$, $\lambda_4 = -0.0802$. Obviously, this result lies in Region 4 of Figure 2. Two different starting points were used. First, we used $\lambda_3 = \lambda_4 = -0.05$, which also lies in Region 4. We then used $\lambda_3 = \lambda_4 = 0.5$, which is in Region 3. The FORTRAN code yielded the tabulated values for both cases, indicating that, in the latter case, the search switched regions.

The particular steps taken in the latter search ($\lambda_3 = \lambda_4 = 0.5$) are presented in Figure 3. The points in Figure 3 represent the termination points of each iteration of Powell's Algorithm, while the segments connecting them represent the path taken from the initial point in Region 3 (top right-hand corner) to the optimal point in Region 4 (lower left-hand corner). As can be seen, the algorithm crosses regions near the (0,0) point, and no iteration termination points lie in the forbidden zone. Figure 4 shows the termination points of all the quadratic approximation searches used as the algorithm approached and crossed the boundary. Again, the segments represent the path taken by the algorithm. As we expect, none of the termination points for the quadratic approximation searches lie in the forbidden region. The algorithm, therefore, *is* obviously capable of jumping directly from one region to the other.

4.4 Constrained Behavior of the Algorithm

Our next concern was with regard to cases where the values for λ_3 and λ_4 that minimize Equation 14 lie within Regions 3 or 4 and close to, or on, one of the limiting boundaries. Our analysis of the algorithm suggests that it might be possible for the algorithm to "miss" the optimal value. An example will demonstrate this. Let us assume that an iteration of Powell's Algorithm concludes at a point near the optimal value (which lies arbitrarily near one of the limiting boundaries), but the point does not cause the algorithm to meet the termination criteria given in Section 4.3. Another set of quadratic approximation searches is therefore performed. Let us further assume that the first such search is correctly in the direction of the optimal point, but also in the direction of the

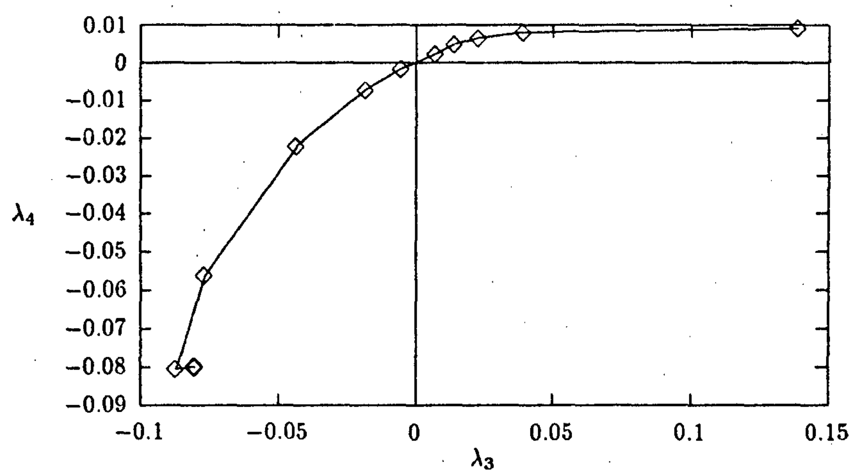

Figure 3. Ending Points for Iterations of Powell's Algorithm: $\alpha_3 = 0.0, \alpha_4 = 6.0$

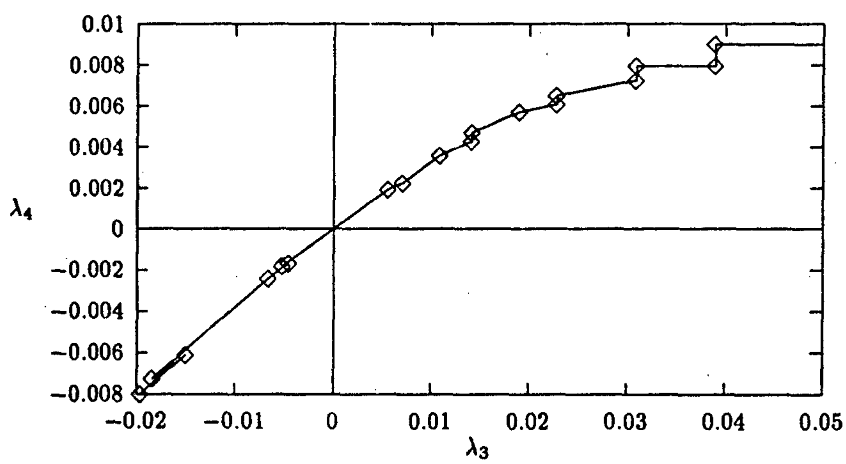

Figure 4. Ending Points for Quadratic Approximation Searches: $\alpha_3 = 0.0, \alpha_4 = 6.0$

"forbidden zone." If the quadratic approximation search encounters a candidate solution either in the "forbidden zone" or at a point where the value of the function is now much larger than before, the search process will switch directions and search *away* from the optimal point. If this were to happen for several successive iterations, it would seem possible for the algorithm's search termination criteria to be met at some point other than the optimal solution—at the "wrong point." Hopefully the additional searches detailed in Steps 2-5 of Table 3 will prevent this from happening, but we cannot be sure.

The choice of starting values for λ_3 and λ_4 play a role in determining the solution at which the algorithm will terminate. Because of this, one possible remedy to the potential problem we have described might be to perform additional searches, using different starting points, as a check. If these agree, then one might have more confidence that this potential problem was not encountered.

Another solution to this problem would be to decrease the step size, q, used in the quadratic approximation searches. By decreasing the interval between successive search points, we can reduce the possibility of the above situation occurring, since we hopefully would be less likely to "skip over" the optimal point or step into the "forbidden zone." This is done at a cost, however. By decreasing the step size, we might increase the number of iterations needed to generate the final solution, which in turn will require more CPU time to solve a particular problem.

Despite these possible solutions, we are still not certain as to what effect the current penalty function has on the search procedure. Does its inclusion force the search algorithm into incorrect directions? How likely is this situation to occur? This is very hard to predict.

As a way of answering these questions, we modified the FORTRAN program of Mykytka and Ramberg [10] to allow any combination of λ_3 and λ_4, *i.e.*, Equation 15 is ignored. By doing this, the algorithm will not be affected by the penalty function, although we must remember that any combination of λ_3 and λ_4 that lies outside of Regions 3 and 4 of Figure 2 corresponds to an invalid pdf.

By comparing the resulting values for the constrained and unconstrained versions, we will be able to determine whether the current penalty function adversely affects the algorithm. If the penalty function is affecting the algorithm, we should see a difference in the final values of λ_3 and λ_4 generated by the two methods. If it has no effect, there should be no difference between the two methods. We choose to examine a number of points that lie on the limiting boundaries (*i.e.*, either λ_3 or λ_4 equals zero), since these are the cases where the penalty function will be a factor. The α_3–α_4 combinations given in Table 4 were chosen using an adapted version of the FORTRAN code of Mykytka and Ramberg [10]. This revised version calculates the resulting α_3–α_4 combination when given specific values of λ_3 and λ_4. A summary of the results is given in Table 4.

Desired	*Desired*	*CONSTRAINED*			*UNCONSTRAINED*		
α_3	α_4	λ_3	λ_4	*Min value*	λ_3	λ_4	*Min value*
-0.5656	2.4000*	0.4999	4.873×10^{-5}	2.604×10^{-8}	0.5000	-4.237×10^{-6}	3.75×10^{-15}
-1.0498	3.6964	0.2500	6.577×10^{-7}	1.877×10^{-9}	0.2500	-6.049×10^{-6}	3.3387×10^{-15}
0.3636	1.8701	1.5000	1.282×10^{-6}	2.671×10^{-10}	1.5000	-2.256×10^{-6}	8.44×10^{-22}
0.4500	2.2000	4.986×10^{-6}	0.5812	1.554×10^{-6}	4.986×10^{-6}	0.5812	1.554×10^{-6}
0.0000	1.8000	1.751×10^{-7}	1.0000	3.223×10^{-13}	6.591×10^{-12}	1.0000	3.94×10^{-20}
1.0498	3.6964	1.576×10^{-6}	0.2500	1.093×10^{-9}	1.576×10^{-6}	0.2500	1.093×10^{-9}
0.1897	1.9009	1.491×10^{-5}	0.7999	1.749×10^{-16}	1.491×10^{-5}	0.7999	1.749×10^{-16}
-0.1969	1.7961	1.849×10^{-5}	1.250	4.647×10^{-17}	1.849×10^{-5}	1.250	4.647×10^{-17}
* *the λ_3 and λ_4 values given for the constrained search yielded $\alpha_4 = 2.4001$.*							

Table 4. Comparison of Unconstrained and Constrained Searches

As Table 4 shows, the search procedure yields almost identical solutions for the two cases. The columns labelled "Min value" represent the sum-of-squared errors calculated using the displayed values of λ_3 and λ_4.

An examination of the table shows that along the $\lambda_3 = 0$ boundary (*i.e.* the last five cases of Table 4), the two searches yield almost identical values for λ_3, λ_4, and the sum-of-squared errors. In fact, the results are identical—except in the case where $\alpha_3 = 0.0$ and $\alpha_4 = 1.8$. In this case, the unconstrained search yields slightly smaller values for λ_3 and the sum-of-squared errors than the constrained search does. Even though this is true, the errors are still sufficiently close to zero, and for all intents and purposes, the two λ_3–λ_4 combinations are the same. We can therefore conclude that the penalty function did not adversely affect the search algorithm for searches along the $\lambda_3 = 0$ boundary.

For all the cases along the $\lambda_4 = 0$ boundary, a smaller final objective function value is possible using the unconstrained search. The sum-of-squared errors for both the unconstrained and constrained cases are approximately equal to zero, but the unconstrained λ_3-λ_4 combinations found in Table 4 are in the "forbidden zone" and therefore do not represent valid pdfs. Although the resulting λ_3-λ_4 combinations from the constrained searches have slightly larger sum-of-squared errors, they still correspond to valid pdfs. For both cases, the resulting values for λ_4 are sufficiently close to zero so that we can safely assume $\lambda_4 = 0$. Since the corresponding λ_3 values are almost identical as well, the penalty function, therefore, does not adversely affect the search algorithm near the $\lambda_4 = 0$ boundary either.

From the cases we have examined, we can see that the use of the current penalty function does not present a problem. Powell's Algorithm behaved consistently—with or without the penalty function. Therefore, we are reasonably confident that we can eliminate the penalty function as a source of error.

4.5 Reparameterization

4.5.1 Derivation. Based on the results of the previous sections of this chapter, it appears that we can reasonably expect Powell's Algorithm to move in the proper fashion when faced with our constraint. However, we can not forget that Powell's Algorithm was expressly designed for unconstrained function optimization. We thus can never be entirely sure that the algorithm will always behave properly when faced with a constrained search region. It, therefore, may be worthwhile to discuss some techniques to avoid using an unconstrained algorithm in this constrained space.

The first, and perhaps simplest method, is to simply ignore the constraint altogether. For the cases shown in Section 4.4, the values of λ_3 and λ_4 calculated using an unconstrained version of Powell's Algorithm were almost identical to those found using the constrained search area. Therefore, we could search with Powell's Algorithm over an unconstrained region, and simply discard any λ_3–λ_4 combination which lies outside of Regions 3 or 4. On the other hand, how do we determine appropriate λ_3 and λ_4 values for these cases? Should we simply conclude that these cases are infeasible? The use of a different starting point might be an option, since the objective function is known to be multi-modal. However, if no appropriate values result from these subsequent searches, we would be forced exclude that particular α_3–α_4 combination from consideration, *i.e.*, we would conclude that that skewness–kurtosis combination can not be modeled by the GLD.

A second approach is to eliminate Powell's Algorithm altogether and replace it with a similar ***constrained*** algorithm. This would involve investigating potential replacements, and then implementing a new routine. Since we are uncertain that any other algorithm would perform any better, this could be a time-consuming, and perhaps unnecessary, process.

A third possibility is to reparameterize the GLD so that its parameters are unconstrained. By doing this, we can eliminate the need for the constraint, Equation 15, thereby changing the minimization problem to one involving an ***unconstrained*** search — a situation for which Powell's Algorithm was expressly designed. This last method seems to hold the most promise for a quick, easy solution to the problem.

There are several different options for reparameterizing the GLD. One of the easiest is the following:

$$\theta_i = \ln(\lambda_i); \;\; \lambda_i = e^{\theta_i}$$

when $\lambda_3, \lambda_4 \geq 0$, and

$$\theta_i = \ln(-\lambda_i); \;\; \lambda_i = -e^{\theta_i}$$

when $\lambda_3, \lambda_4 < 0$.

By changing the variables in this manner, we create an unconstrained search area. Since $e^x \geq 0$ for all x, we are assured that the original parameters, λ_3 and λ_4, will always have the same sign, and we therefore do not need the $\lambda_3 \cdot \lambda_4 \geq 0$ constraint.

It should be noted that this reparameterized version will not be able to switch from one valid region to the other, as the original sometimes did. This is not a big sacrifice, since we know that the original usually did not switch regions, and therefore could not rely on such behavior. If the new search algorithm fails to achieve an acceptable solution when given a starting value in one region, we need simply rerun the FORTRAN program using a starting point in the opposite region.

The reparameterization is accomplished by altering the FORTRAN code used by Mykytka and Ramberg [10]. The main change involves switching the starting values for λ_3 and λ_4 to the new, reparameterized form. After the initial λ values are input by the user, they are calculated in θ form before being sent to the FORTRAN subroutine which performs Powell's Algorithm.

The minimization problem therefore becomes:

$$min\ Z = f(\theta_3, \theta_4) = [(\alpha_3(\theta_3, \theta_4) - \alpha_3)^2 + (\alpha_4(\theta_3, \theta_4) - \alpha_4)^2],$$

which would involve recalculating the equations for α_3 and α_4, Equations 8 and 9, using the new theta parameters—a time-consuming process. An alternative to this is to perform the Powell's Algorithm search procedure in the unconstrained θ space, but evaluate the Z-values at each point in the iteration using the λ values that correspond to the calculated θs. This is easy to accomplish, since the FORTRAN program utilizes a separate subroutine to calculate the values of α_3 and α_4 for a given λ_3–λ_4 combination. We need simply change the θs back into their respective λs upon entering this subroutine to determine the values of α_3 and α_4. We therefore use the new theta parameters only in the subroutine which performs Powell's Algorithm, while the rest of the program remains more or less the same.

In order to prove the validity of this technique, the method was checked with $\alpha_3 - \alpha_4$ combinations documented in the tables of Ramberg *et al* [13]. The new method did achieve the same values of λ_3 and λ_4, although sometimes two separate searches were necessary in the reparameterized case[3]. Since the results of the two methods were identical, we can attempt to expand to values beyond the tabulated range.

4.5.2 Results. This expansion was accomplished by first selecting a starting α_3–α_4 combination from the tables of [13]. By choosing a α_3–α_4 combination whose λ_3 and λ_4 values are already known, we assure the revised FORTRAN program is working properly and give ourselves a basis for comparison between the original and reparameterized versions. After the starting α_3 and α_4 values are chosen, we use the reparameterized FORTRAN program to find the appropriate λ_3 and λ_4 parameters. If the values of λ_3 and λ_4 yield a sum–of–squared error below our acceptable tolerance (approximately 0.0001), the value of α_4 is decremented by 0.2 and the FORTRAN program is run again. When the sum–of–squared errors exceeds our tolerance, we chose a different value of α_3 and restart the process.

Several different values of α_3 were chosen For the α_3–α_4 combinations listed in the tables of [13], the reparameterized version yielded the same λ_3–λ_4 combinations. However, the reparameterized version was not able to expand beyond the tabulated values of Ramberg *et al.*

4.6 Conclusions

This chapter has looked extensively at our particular implementation of Powell's Algorithm and our current penalty function as sources of possible limitations on the range of α_3–α_4 combinations that can be modeled using the GLD. Section 4.1 examined the inner workings of Powell's

[3] This was usually because the initial guess lay in the opposite region as the final result.

Algorithm to provide a better understanding for how the search procedure is actually performed. It also discussed the potential problems the penalty function might cause when trying to find λ_3–λ_4 solutions located near the limiting boundaries.

Section 4.4 showed that the penalty function does not adversely affect the search algorithm when searching for values located near the boundaries. Table 4 shows that the results between the constrained and unconstrained searches were consistent along the $\lambda_3 = 0$ boundary, and that only slight differences existed between the two sets of searches along the $\lambda_4 = 0$ boundary. Even though the unconstrained search yielded lower sum–of–squared errors along the $\lambda_4 = 0$ boundary than the constrained case, the resulting λ_3–λ_4 combinations were in the "forbidden zone." The results from the constrained case along the $\lambda_4 = 0$ boundary were inside the feasible region, and although the sum–of–squared errors were slightly larger than the unconstrained case, they were still well below acceptable tolerances.

Section 4.5 described alternatives to the current situation. In particular, it described the results of a reparameterization of the λ parameters to parameters that are unrestricted in sign. These unrestricted parameters could not improve the range of α_3–α_4 combinations that can be modeled using the GLD.

These results, when taken together, show that the present implementation of Powell's Algorithm does not significantly limit the GLD. Even in cases where the final λ_3 and λ_4 values lie on (or near) the limiting boundaries, the algorithm was not affected by the use of a penalty function. The algorithm also performed just as well in an unconstrained search as it did in the constrained case. We therefore can safely move on to investigate other issues.

V. *Graphical Analysis of the GLD*

5.1 *Introduction*

Mykytka used the tables in [9:87-117][1] to identify the tentative range of α_3–α_4 combinations that are possible to model using the GLD [9:22], which is depicted here in Figure 1. He also presents a contour map of α_3 and α_4 values for combinations of λ_3 and λ_4 each restricted to the range $(0, 1)$ We reprint his contour map ([9:39]) here as Figure 5.

The tables used to generate Figures 1 and 5 are not necessarily complete. As we mentioned in Chapter II, the smallest value of α_4 listed in each table for a given α_3 does not necessarily represent the minimum value of α_4 that is possible for that value of α_3 using the GLD. It is merely the smallest value found for which the optimal solution of the objective function, Equation 14, was approximately zero. It may, therefore, be possible that values of the lambda parameters for values of α_4 smaller than those shown in the table do indeed exist, but for unknown reasons, the FORTRAN program was unable to find them.

The research outlined in this chapter attempts to extend the boundary presented in Figure 1 using a different approach to the problem. As a result of this effort, we will discover an interesting simplification to the equations for α_3 and α_4 that holds under certain conditions.

5.2 *Background*

An examination of Mykytka's tables [9:87-117] shows that the λ_3 and λ_4 values tend to repeat certain behaviors. A representative table is reprinted here in Table 5. As Table 5 shows, for a given value of α_3, as the value of α_4 is decreased (moving from the bottom to the top of a particular table), the corresponding λ_3 value starts negative, and increases towards zero. The value of λ_3 then takes on a positive value, increases for a while, and then decreases towards zero. The corresponding

[1]These tables can also be found in Ramberg *et al* [13].

λ_4 values also start as negatives, increase towards zero, and become positive, but then, instead of decreasing towards zero, the λ_4 values continue to increase.

Since this behavior is repeated for most of the tables of α_3 and α_4 found in [13], we hypothesize that, for a given value of α_3, the minimum value of α_4 possible using the GLD will occur when $\lambda_3 = 0$.

Further evidence of this can be seen in Figure 5. This figure shows a subset of the values of α_3 and α_4 that result from given combinations of λ_3–λ_4. We can see that for a given value of α_3, the minimum α_4 occurs when $\lambda_3 = 0$. On the other hand, Figure 5 only shows a small (although important) subset of the possible values of λ_3 and λ_4, *i.e.* those between zero and 1, so we might witness different behavior for value of λ_3 and λ_4 outside of this range.

As a first step in attempting to prove this hypothesis, we will attempt to expand Figure 5 by exploring regions outside of the range $0 \leq \lambda_3, \lambda_4 \leq 1$. By extensively examining these regions, we will hopefully find at least a graphical verification of our assumption.

It should be noted here that as we extend the range of points covered, we will most likely encounter "less-useful" (or, at least, less intuitively-appealing) pdfs. The shaded region of Figure 1, which signifies the range of α_3–α_4 combinations that can be modeled using the GLD, already contains most of the commonly-used pdfs. The only group of pdfs not fully covered are those that can be represented by the U– and J–shaped beta distributions. An examination of the GLD's pdf, Equation 2.1, shows that as λ_3 is increased above one, the resulting pdf will have a non-zero (i.e. positive) value at its left endpoint (*i.e.*, $f(x) > 0$ at the lower limit of the distribution's range). Similarly, as λ_4 is increased beyond one, the resulting function will have a non-zero value at its right endpoint. Since most of the U– and J–shaped beta distributions pdf's exhibit similar behavior, we suspect that by extending the range beyond $0 \leq \lambda_3, \lambda_4 \leq 1$, we will encounter α_3–α_4 combinations characteristic of these distributions.

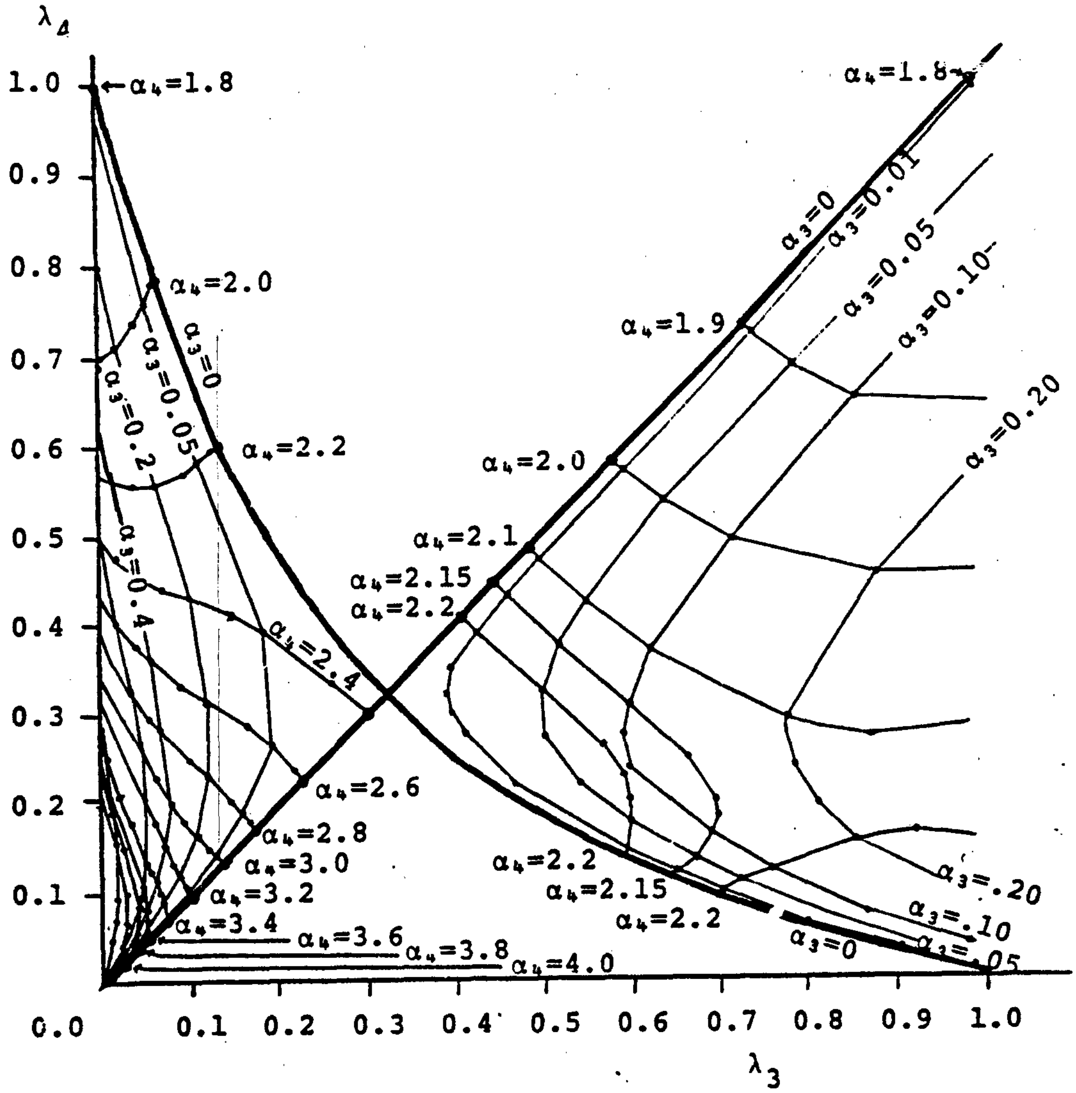

Figure 5. Selected Contours of $\alpha_3(\lambda_3,\lambda_4)$ and $\alpha_4(\lambda_3,\lambda_4)$ in Region 3

A3 = 0.30

A4	LAM 1	LAM 2	LAM 3	LAM 4
2.0	-1.550	.2660	.0000	.7020
2.2	-1.164	.2755	.0380	.5556
2.4	-.871	.2733	.0695	.4348
2.6	-.642	.2586	.0911	.3324
2.8	-.474	.2323	.0983	.2495
3.0	-.362	.1991	.0925	.1859
3.2	-.288	.1641	.0796	.1377
3.4	-.239	.1298	.0640	.1003
3.6	-.204	.0973	.0481	.0704
3.8	-.179	.0671	.0330	.0460
4.0	-.160	.0389	.0190	.0255
4.2	-.144	.0127	.6175+	.8035+
4.3	-.138	.0789+	.0380+	.0489+
4.4	-.131	-.0116	-.5554+	-.7057+
4.5	-.129	-.0231	-.0110	-.0139
4.6	-.121	-.0343	-.0163	-.0203
4.8	-.113	-.0554	-.0260	-.0319
5.0	-.105	-.0752	-.0350	-.0423
5.2	-.100	-.0939	-.0432	-.0517
5.4	-.094	-.1114	-.0508	-.0601
5.6	-.089	-.1279	-.0578	-.0678
5.8	-.085	-.1435	-.0643	-.0748
6.0	-.081	-.1582	-.0703	-.0812
6.2	-.078	-.1722	-.0759	-.0872
6.4	-.075	-.1854	-.0811	-.0927
6.6	-.072	-.1979	-.0860	-.0977
6.8	-.069	-.2100	-.0906	-.1025
7.0	-.067	-.2214	-.0949	-.1069
7.2	-.065	-.2325	-.0990	-.1111
7.4	-.063	-.2427	-.1028	-.1149
7.6	-.061	-.2528	-.1064	-.1186
7.8	-.060	-.2623	-.1098	-.1220
8.0	-.058	-.2716	-.1131	-.1253
8.2	-.056	-.2805	-.1162	-.1284
8.4	-.055	-.2889	-.1191	-.1313
8.6	-.054	-.2971	-.1219	-.1341
8.8	-.053	-.3050	-.1246	-.1367
9.0	-.052	-.3125	-.1271	-.1392
9.2	-.051	-.3197	-.1295	-.1416

Table 5. Sample Table of α_3–α_4 Combinations: $\alpha_3 = 0.3, 2.0 \geq \alpha_4 \geq 9.2$

5.3 A Simplification of the Equations for α_3 and α_4

A nice sidelight to our hypothesis is that when $\lambda_3 = 0$, the equations for α_3 and α_4 (Equations 8 and 9) can be written in a much simpler form. We make use of the fact that

$$\beta(1,a) = \beta(a,1) = \frac{\Gamma(1)\Gamma(a)}{\Gamma(a+1)} = \frac{1}{a},$$

so that when $\lambda_3 = 0$, the equations for the intermediate parameters A, B,C, and D (given by Equations 10 through 13) can be reduced to:

$$\begin{aligned} A &= 1 - \frac{1}{1+\lambda_4} \\ &= \frac{\lambda_4}{1+\lambda_4} \end{aligned}$$

$$\begin{aligned} B &= 1 - 2\beta(1, 1+\lambda_4) + \frac{1}{1+2\lambda_4} \\ &= \frac{2\lambda_4^2}{(1+\lambda_4)(1+2\lambda_4)} \end{aligned}$$

$$\begin{aligned} C &= 1 - 3\beta(1, 1+\lambda_4) + 3\beta(1, 1+2\lambda_4) - \frac{1}{1+3\lambda_4} \\ &= \frac{6\lambda_4^3}{(1+\lambda_4)(1+2\lambda_4)(1+3\lambda_4)} \end{aligned}$$

$$\begin{aligned} D &= 1 - 4\beta(1, 1+\lambda_4) + 6\beta(1, 1+2\lambda_4) - 4\beta(1, 1+3\lambda_4) + \frac{1}{1+4\lambda_4} \\ &= \frac{24\lambda_4^4}{(1+\lambda_4)(1+2\lambda_4)(1+3\lambda_4)(1+4\lambda_4)} \end{aligned}$$

The formulas for α_3 and α_4 become:

$$\alpha_3 = \frac{2(1+\lambda_4)\sqrt{1+2\lambda_4}}{1+3\lambda_4}$$

$$\alpha_4 = \frac{2(1+2\lambda_4)(5+3\lambda_4+16\lambda_4^2+12\lambda_4^3)}{(1+3\lambda_4)(1+4\lambda_4)}$$

Note that the equations for α_3 and α_4 are now functions of a single variable, λ_4. Since we have two equations in a single unknown, we should be able to solve for λ_4 when given a specific α_3 value. If our hypothesis is valid, we can then use this value of λ_4 to find the minimum possible value of α_4 for the given value of α_3 and we can therefore establish a lower limit for the kurtosis.

Because of the symmetry of λ_3 and λ_4 in the equations for A, B, C, and D, we also found that, for the case where $\lambda_4 = 0$, the equation for α_4 remains the same (with, of course, λ_3 replacing λ_4), while that for α_3 becomes the negative of the $\lambda_3 = 0$ case, *i.e.:*

$$\alpha_3 = -\frac{2(1+\lambda_3)\sqrt{1+2\lambda_3}}{1+3\lambda_3}$$

$$\alpha_4 = \frac{2(1+2\lambda_3)(5+3\lambda_3+16\lambda_3^2+12\lambda_3^3)}{(1+3\lambda_3)(1+4\lambda_3)}$$

5.4 Procedure

A FORTRAN program was written which calculates the α_3–α_4 values which correspond to a specified combination of λ_3 and λ_4. The program was adapted from the subroutine in the FORTRAN code of Mykytka and Ramberg [10] that accomplished the same task. This program will be a valuable tool for subsequent analysis, since we will be able to restrict our attention to only valid α_3–α_4 combinations. Instead of choosing a combination of skewness and kurtosis and running the original FORTRAN program (as Mykytka [9] did)—hoping the GLD can find a λ_3–λ_4 combination to replicate them—we will do something much easier. By setting the values for λ_3 and λ_4 first, we know that the resulting α_3 and α_4 values must be within the GLD's range (as long as only λ_3–λ_4 combinations in Regions 3 and 4 are used). By checking a large enough subset of

λ_3–λ_4 combinations, we hopefully will be able to get a better idea of the range of possible α_3–α_4 combinations that can be modeled using the GLD.

5.5 *Results*

We originally started by re-examining the range $0 \le \lambda_3, \lambda_4 \le 1$. Since an infinite number of different λ_3–λ_4 combinations are possible over this range, we obviously can only look at a small fraction of the total number of values. Figure 6 shows the combinations of α_3 and α_4 that result when λ_3 and λ_4 are each iterated over this range in steps of 0.01 each.[2] We can note several interesting details from Figure 6. First of all, the α_3–α_4 combinations are symmetric around the $\alpha_3 = 0$ axis. This is expected, since Equations 8 and 9 show that when the values of λ_3 and λ_4 are interchanged, the resulting distribution has the same α_4 value, but has a α_3 value that is the negative of the original. Therefore, we should expect this type of symmetric behavior.

Figure 6 also shows that there is an apparent lower limit on the possible α_4 values, but that the minimum possible α_4 value for any particular case depends on the associated value of α_3. For example, we can see that a minimum α_4 value of approximately 1.8 is observed, but only when $\alpha_3 = 0$. When $\alpha_3 = 1$, the minimum observed value of α_4 is only 3.5.

Figure 7 shows the α_3–α_4 combinations that result when we set $\lambda_3 = 0$ and iterate λ_4 over the range in 0.01 intervals. Comparing Figures 6 and 7 shows that the minimum α_4 value for a given value of α_3 ($\alpha_3 \ge 0$) does indeed occur when $\lambda_3 = 0$. Figure 8 shows the results when we set $\lambda_4 = 0$ and vary the value of λ_3 over the same range. Obviously, when $\alpha_3 \le 0$, the minimum value of α_4 (for a given value of α_3) occurs when $\lambda_4 = 0$, as we would expect. The two figures also show that when $\alpha_3 = 0$, setting either λ_3 or λ_4 to zero results in the same minimum value of α_4.

So, over our limited range of values, it seems that our hypothesis is valid. However, what happens when we iterate over the same range in intervals smaller than 0.01? Does our hypothesis

[2] We disregard the combination $\lambda_3 = \lambda_4 = 0$ for this case, as well as all future cases, since it yields a GLD with an invalid pdf.

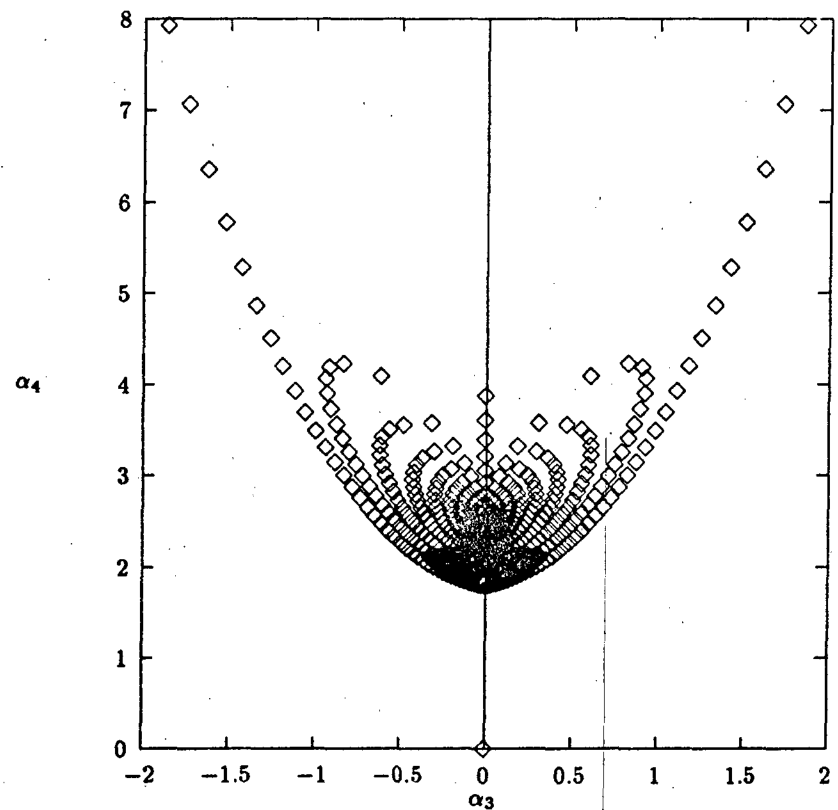

Figure 6. α_3–α_4 Combinations for $0 \leq \lambda_3, \lambda_4 \leq 1$

still hold? To answer this question, we studied cases where the step size was set as low as 0.0001. These additional points, when added to Figure 6, merely "filled in" the range above our previous limits. Setting λ_3 (or λ_4, depending on the range of interest) equal to zero still gave the smallest α_4 value for a given value of α_3.

Our next step was to attempt to expand our hypothesis to regions beyond the range $0 \leq \lambda_3, \lambda_4 \leq 1$. Although we could not inspect every possible λ_3–λ_4 combination, we tried to examine typical cases over the entire range of possible values.

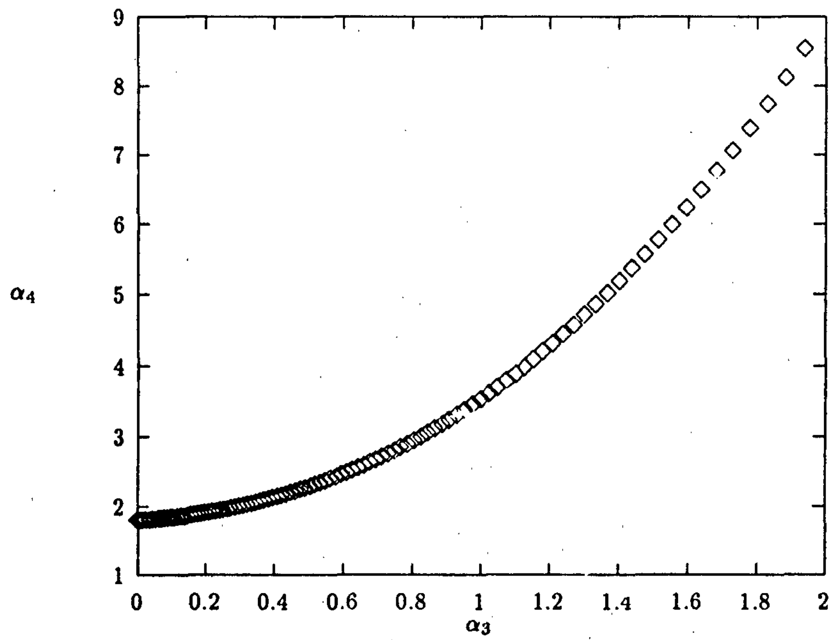

Figure 7. α_3–α_4 Combinations for $\lambda_3 = 0, 0 < \lambda_4 \leq 1$

To do this, we first needed to determine maximum values for λ_3 and λ_4. The lower bound on λ_3 and λ_4 had already been given by Equation 5, $min(\lambda_3, \lambda_4) > -\frac{1}{4}$. However, there were no restrictions on the largest possible values these parameters could have. Through trial and error, a working upper bound of $\lambda_3, \lambda_4 \leq 40$ was proposed. Although specific combinations of λ_3 and λ_4 exist where one of the parameters has a value greater than 40, we wished to study those combinations where we can vary both variables over their entire ranges without encountering any undefined α_3–α_4 combinations. We do not expect that we are dismissing a large number of λ_3–λ_4 combinations by using this assumption, and it gives us an upper bound with which to work.

Figure 9 shows the resulting α_3–α_4 combinations when λ_3 and λ_4 were varied over the range $0 < \lambda_3, \lambda_4 \leq 40$ using a step size of 1.0. Due to limitations in the capacity of our plotting program, this larger step size was needed to plot this entire range. We expect that as we saw before, as

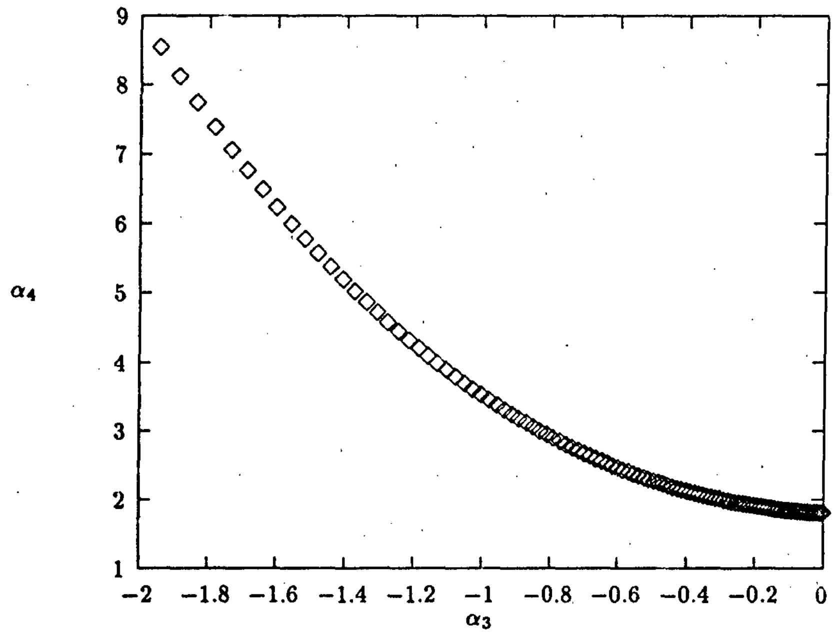

Figure 8. α_3–α_4 Combinations for $0 < \lambda_3 \leq 1, \lambda_4 = 0$

the step size is decreased (thereby increasing the total number of λ_3–λ_4 combinations plotted), the new points will be located in positions that will cause the open area above the parabolic curve to be "filled in."

Figures 10 and 11 show the resulting α_3–α_4 combinations when we set one of the parameters equal to zero and varied the other over the range from zero to 40, using a step size of 0.01. Figure 10 set $\lambda_3 = 0$ and varied λ_4 over this range, while Figure 11 set $\lambda_4 = 0$ and varied λ_3.

By comparing Figures 10 and 11 to the complete set of points in Figure 9, we can see that these two subsets of values do indeed encompass the minimum α_4 values for all of the α_3 values. Figure 12 shows the results of Figures 10 and 11 together on the same set of axes.

By examining these two subsets of values in Figure 12 together, it is obvious that over α_3's negative range, setting $\lambda_3 = 0$ produces the corresponding minimum α_4 value, while over α_3's

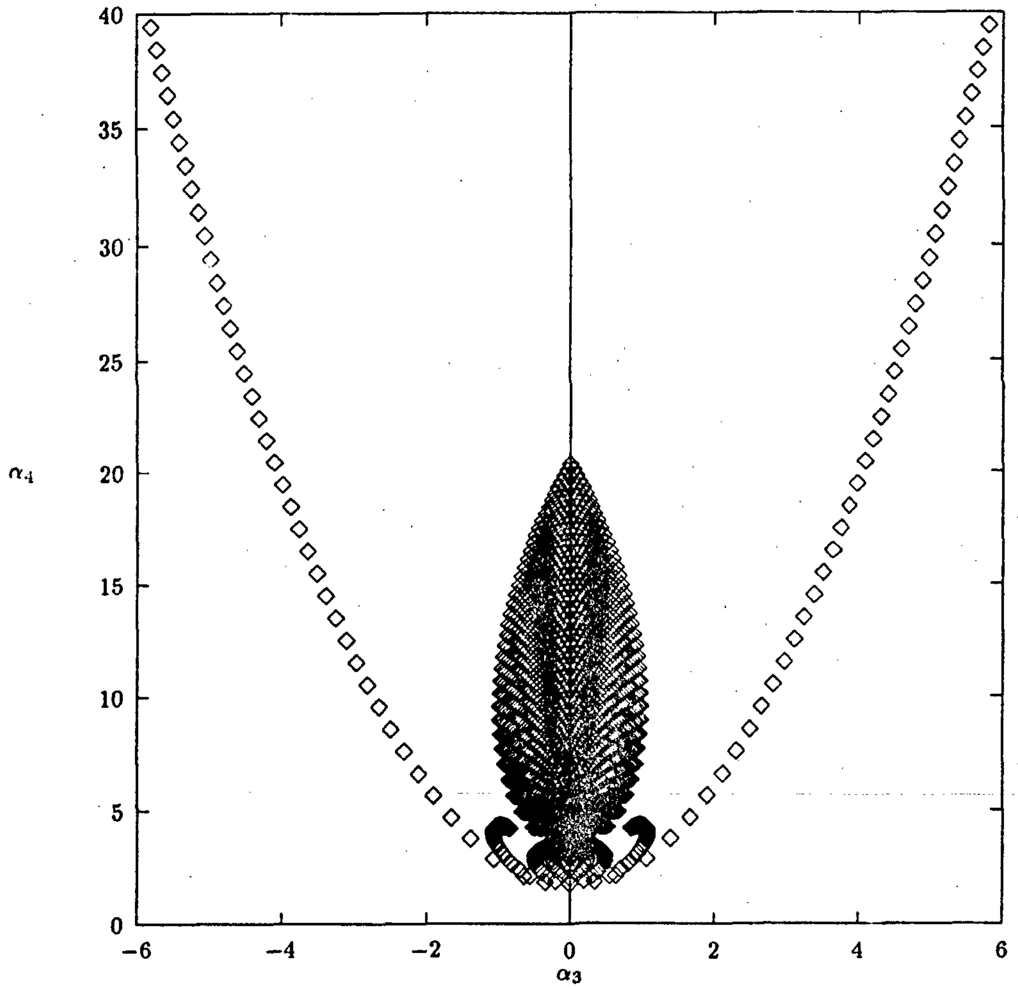

Figure 9. α_3–α_4 Combinations for $0 < \lambda_3, \lambda_4 \leq 40$

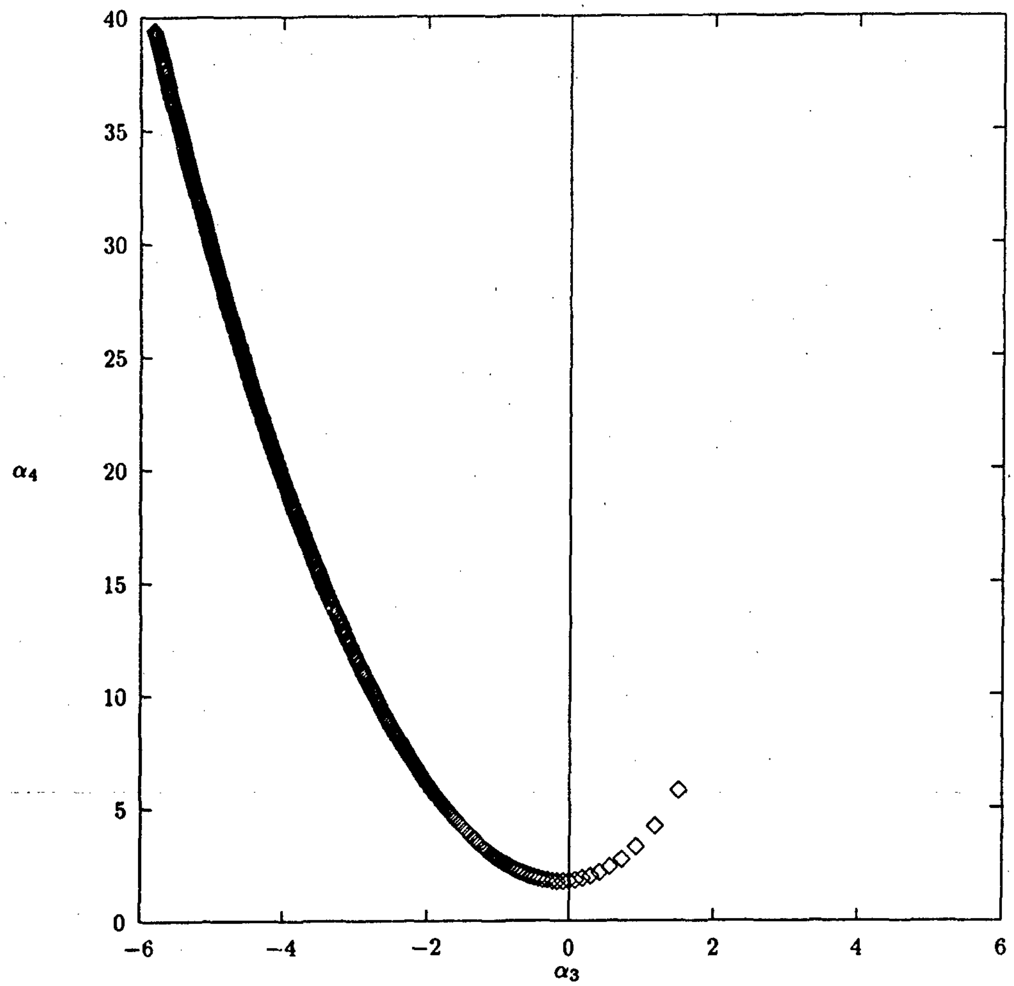

Figure 10. α_3–α_4 Combinations for $\lambda_3 = 0, 0 < \lambda_4 \leq 40$

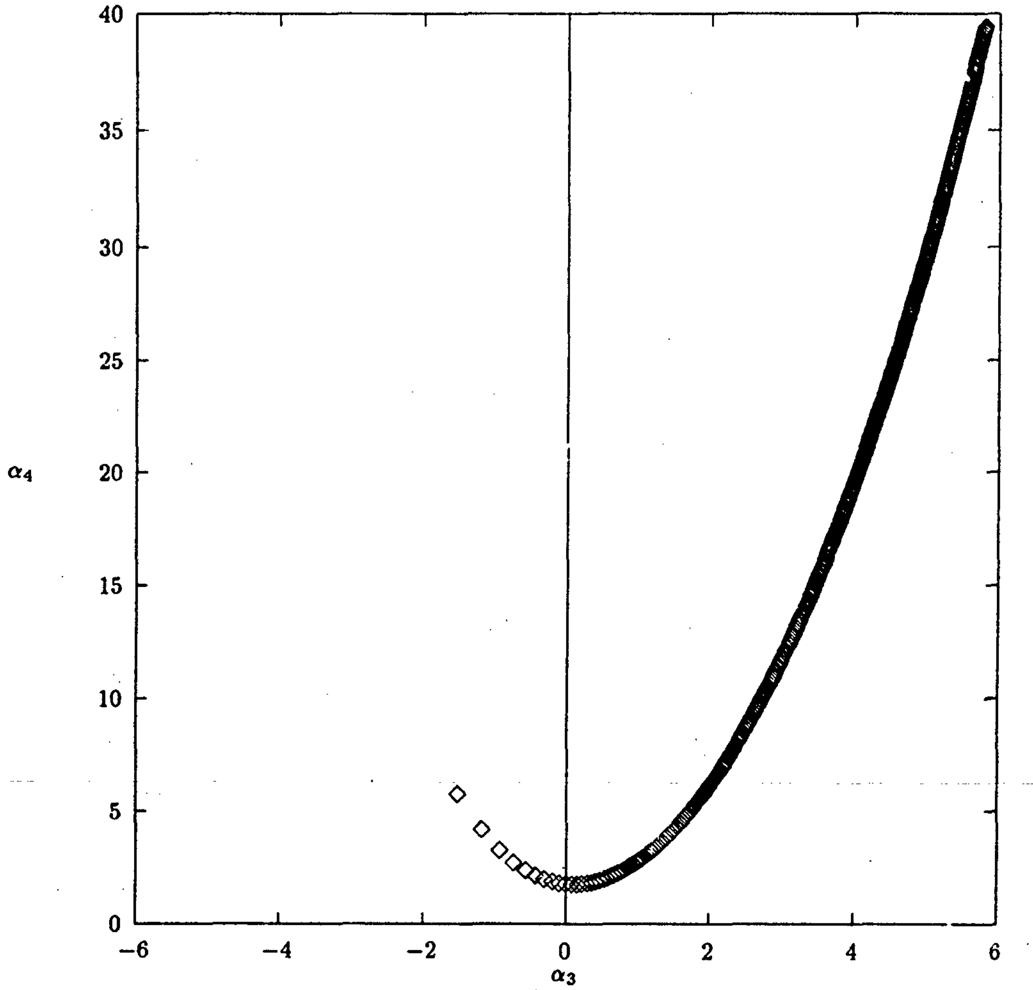

Figure 11. α_3-α_4 Combinations for $\lambda_4 = 0, 0 < \lambda_3 \leq 40$

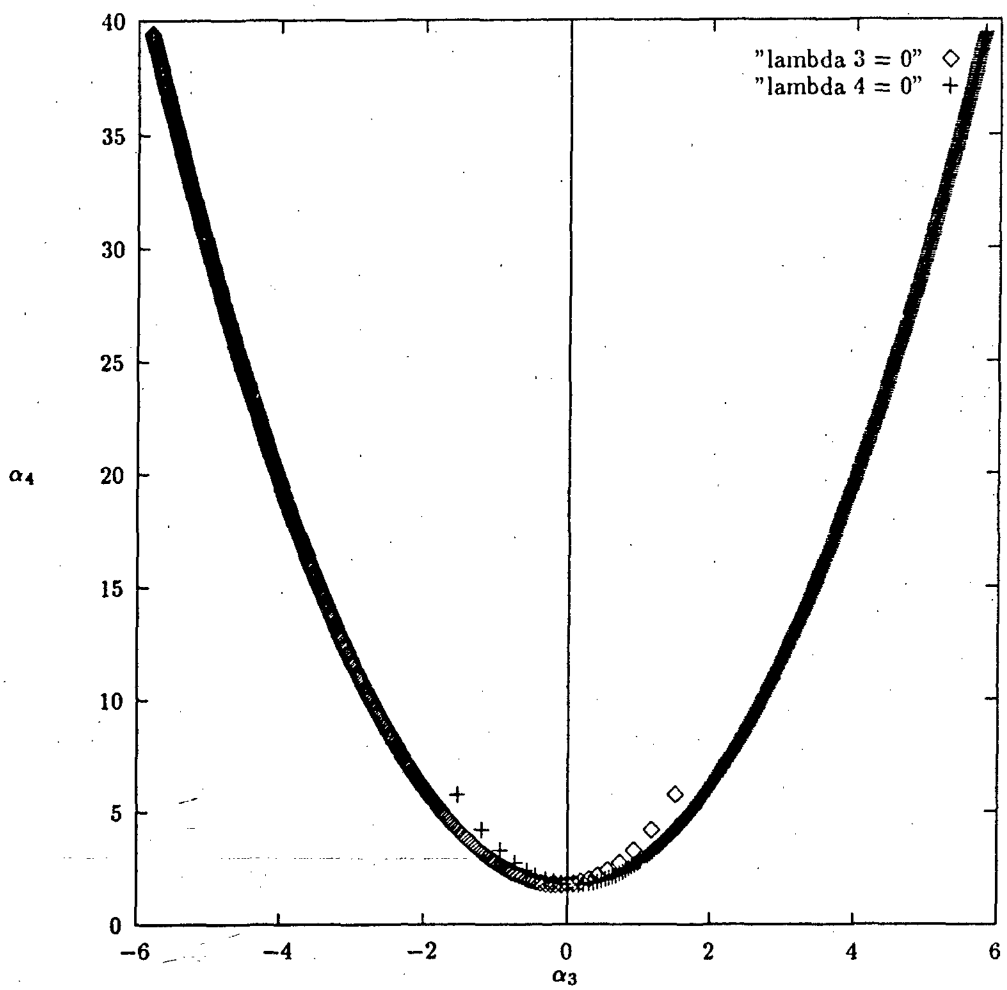

Figure 12. α_3–α_4 Combinations for $\lambda_4 = 0$ and $\lambda_3 = 0$

positive range, setting $\lambda_4 = 0$ produces the minimum α_4 value—the "opposite" of our results from the earlier case, where λ_3 and λ_4 were restricted to values between zero and one.

Figure 13 shows the area where the two sets of points cross in more detail. Since Figure 12 includes a subset of the range of values shown in Figures 7 and 8, we can see where these earlier points lie in our extended range. Obviously, the points in Figure 7 correspond to the $\lambda_3 = 0$ points found above the $\lambda_4 = 0$ curve in Figure 12 (in the range where $\alpha_3 \geq 0$). Similarly, the points found in Figure 8 correspond to the $\lambda 4 = 0$ points found above the $\lambda 3 = 0$ points in Figure 12 ($\alpha_3 \leq 0$).

An examination of the λ_3–λ_4 combinations used to generate Figure 13 shows that these "upper" points correspond to the cases where the non-zero parameter had a value less than 1.0. The two points shown at $\alpha_3 = 0$ in Figure 12 correspond to the pairs $\lambda_3 = 0, \lambda_4 = 1$ and $\lambda_3 = 1, \lambda_4 = 0$. As the non-zero parameter is increased above 1.0, the "lower" part of the respective curve is formed. As it is decreased below 1.0, the "upper" part of the curve is generated.

It therefore appears that the sets of values given in Figures 7 and 8 are *not* the minimum α_4 values that are possible using the GLD. By setting either λ_3 or λ_4 equal to zero (depending on whether a negative or positive value of α_3, respectively, is desired) and allowing the other to take on values above one, we can attain smaller values of α_4 for the same values of α_3.

To this point in our analysis, we have only examined λ_3–λ_4 combinations in Region 3. However, there are also valid λ_3–λ_4 combinations in Region 4. What values of α_3 and α_4 do these combinations yield?

Figure 14 shows the α_3–α_4 combinations that result when λ_3 and λ_4 are varied over the range $-.25 < \lambda_3, \lambda_4 < 0$ using a step size of 0.01. We can see that although the range of α_3 is similar to the previous cases, the resulting values of α_4 are *much* larger. Figure 15 shows only the lower portion of this range, using a smaller step size (thereby showing more points). Obviously, none of these λ_3–λ_4 combinations yield a α_4 value lower than those we have already found, but instead "fill in" the upper range of α_3–α_4 combinations not covered by the λ_3–λ_4 combinations of Region 4.

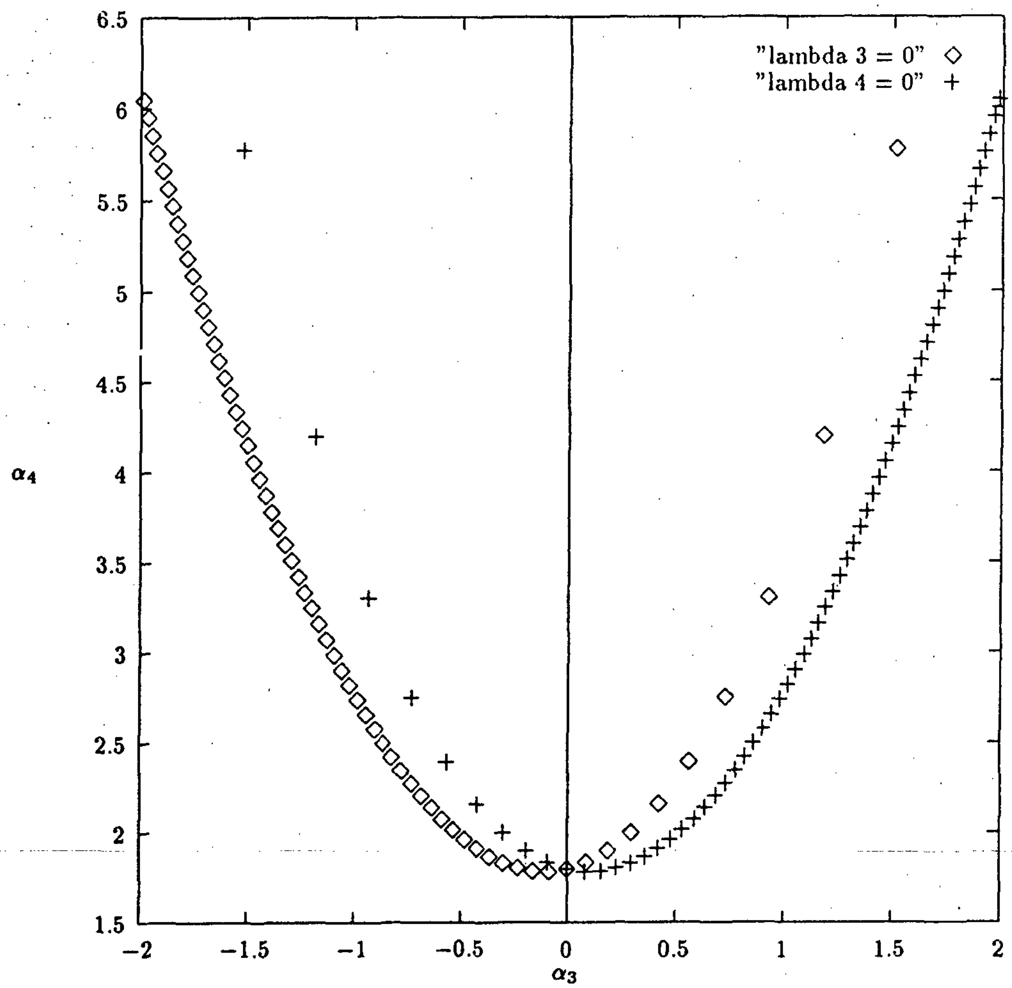

Figure 13. Expansion of Crossing Region

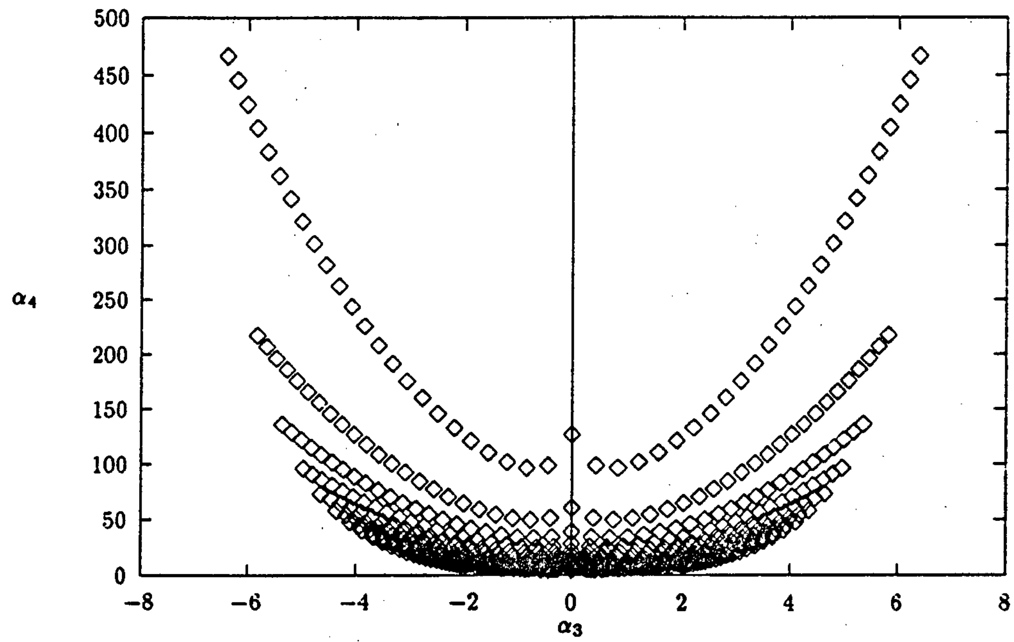

Figure 14. α_3–α_4 Combinations for $-.25 < \lambda_3, \lambda_4 < 0$

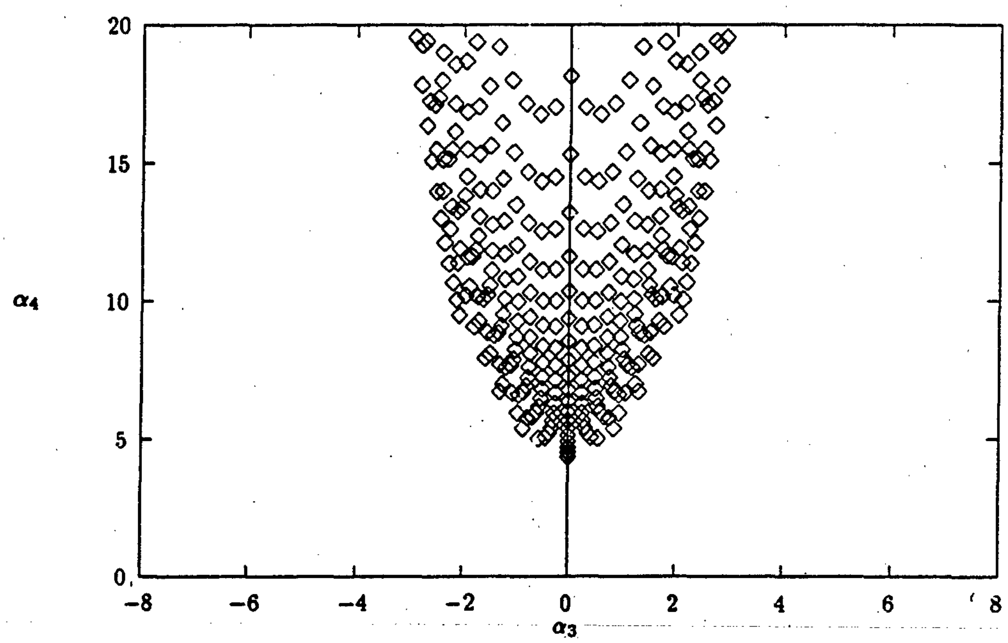

Figure 15. Expansion of Lower Range

5.6 *Conclusions*

Our original hypothesis was that for a given value of α_3, the minimum possible value of α_4 that the GLD can produce occurs when $\lambda_3 = 0$. When we examined the range $0 < \lambda_3, \lambda_4 \leq 1$, it appeared that this hypothesis was at least partly true. However, when we examined λ_3–λ_4 combinations outside of this range, we noticed that a lower value of α_4 could be found for the same α_3.

We must therefore revise our original hypothesis. A more reasonable one seems to be that for a given non–negative value of α_3, the minimum α_4 value that the GLD can produce occurs when $\lambda_4 = 0$ and $\lambda_3 \geq 1$. If α_3 is negative, the minimum α_4 value occurs when $\lambda_3 = 0$ and $\lambda_4 \geq 1$.

If we look at the sample table given in Figure 5, we can most likely "improve" on its smallest α_4 value. We expect that as we decrease the value of α_4 below the smallest tabulated value, the corresponding λ_3 values should decrease to zero, and then rise again to values above 1.0 at its true minimum. The corresponding values of λ_4 should decreases to zero, and become zero for the minimum possible α_4. This will definitely increase the GLD's present range.

Although we have only looked at only a small subset of the λ_3–λ_4 combinations, we can be confident of our results. It was mentioned previously that we needed to use a fairly large step size in some of our analysis. When this step size was decreased, we did not witness any unexpected behavior. Instead, these additional point simply fill in the α_3–α_4 space above our minimum α_4 curve. If we were to use an infinitesimally small increment between λ_3 and λ_4 values, we expect to see a solid region above the minimum α_4 line, as shown by Figure 16.

As a final result, Figure 17 shows both the original coverage region of the GLD (from Figure 1) and our newly–covered α_3–α_4 combinations (denoted by the darker shaded region). Although we can not reach the boundary of the Impossible Area, we can indeed cover a larger portion.

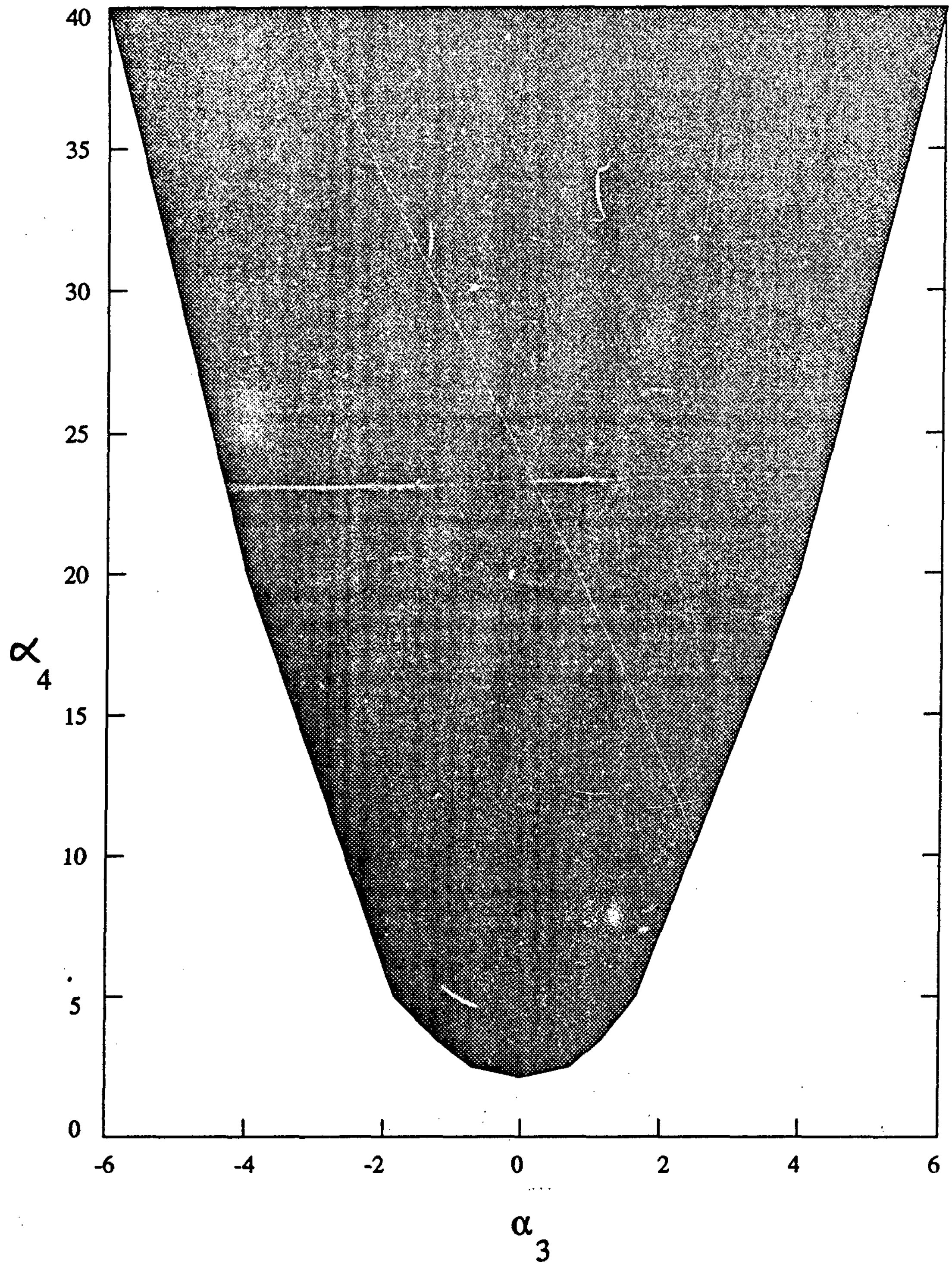

Figure 16. Range of α_3–α_4 Combinations Possible Using the GLD

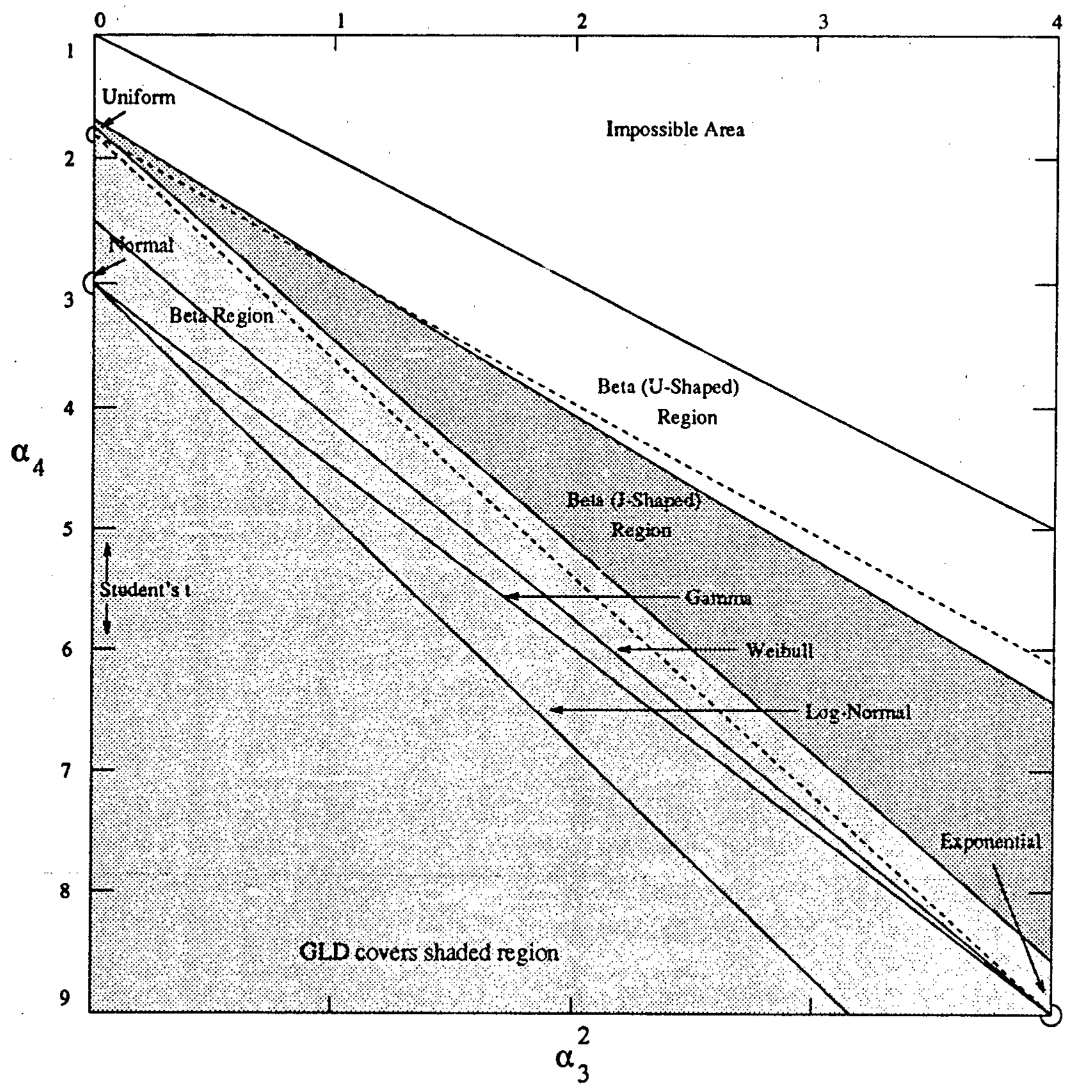

Figure 17. Expanded Coverage Range of the GLD

VI. *Examination of Uncovered α_3–α_4 Range*

The results of Chapter V showed that we can extend the present range of α_3–α_4 combinations that the GLD can model. However, as Figure 17 shows, we are still not able to to cover the entire spectrum of possible α_3–α_4 combinations. What do these "uncovered" distributions look like? Is it even worthwhile to try to include them? This chapter will try to answer these questions by looking at a few specific cases that lie outside of the GLD's range.

From Figure 17, we can see that the α_3–α_4 combinations not covered by the GLD can be modeled by U– or J–shaped beta distributions. Bury [1] gives the following equations for the skewness and kurtosis (what he calls the first and second shape factors) of the two–parameter beta distribution

$$\alpha_3 = \frac{2(b-a)}{a+b+2}\sqrt{\frac{a+b+1}{ab}} \tag{16}$$

(17)

$$\alpha_4 = \frac{3(a+b+1)[2(a-b)^2+ab(a+b+2)]}{ab(a+b+2)(a+b+3)}, \tag{18}$$

where a and b represent the two beta distribution parameters. Since we know what values of skewness and kurtosis we desire, we neea simply solve these two equations to find the appropriate values for a and b. We can then plot the resulting distribution functions using the equation for the beta pdf,

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} \quad 0 < x < 1. \tag{19}$$

We will look at the four α_3–α_4 combinations shown (along with their respective beta parameters, a and b) in Table 6. By looking at Figure 17, we can see that these cases correspond to two U–shaped and two J–shaped beta distributions that can not be duplicated using the GLD. Although we are only looking at a small number of cases, we should at least get a feel for the

Case	α_3	α_4	a	b
1	$\sqrt{0.5}$	2	0.28	0.578
2	$\sqrt{3}$	5	0.19	1.01
3	2	6.25	0.179	1.185
4	$\sqrt{3}$	4.3	0.048	0.233

Table 6. Four Uncovered α_3–α_4 Combinations

type of shapes we can expect from other pdfs that fall into these regions. Figures 18 through 21 show the four resulting pdfs. Note that these four pdfs seem to be "extreme" examples of beta distributions—interesting, but perhaps not very useful.

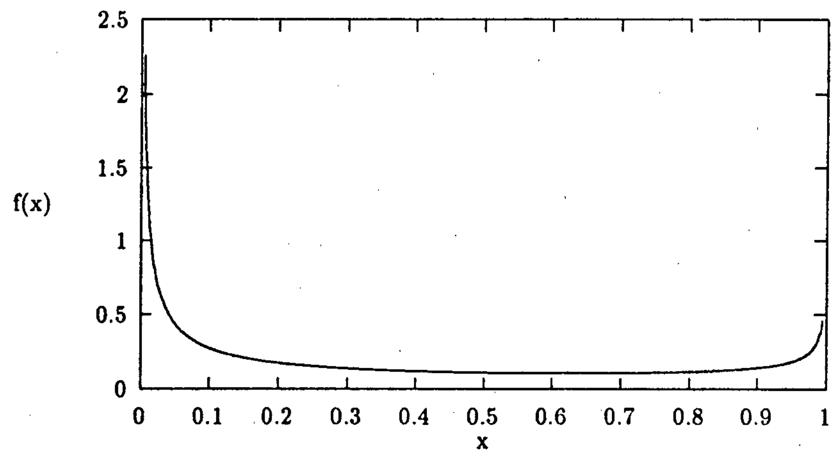

Figure 18. Distribution Function for Case 1

To see this point, recall that one of our reasons for wanting to expand the GLD's range was to make it more useful as a simulation tool. By expanding its range, we can model a wider range of α_3–α_4 combinations, and hence a wider range of potential empirical data sets. However, we must also look at the situation realistically. In ordinary simulations, we do not expect to see many empirical data sets that resemble the pdfs shown in Figures 18 through 21.

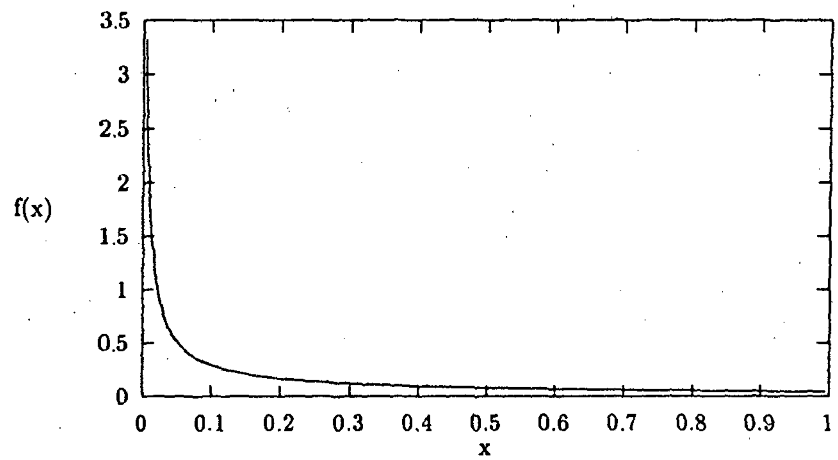

Figure 19. Distribution Function for Case 2

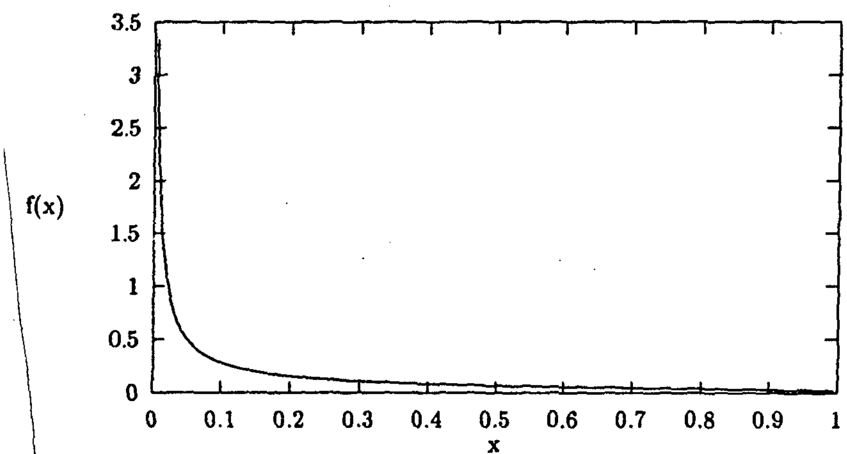

Figure 20. Distribution Function for Case 3

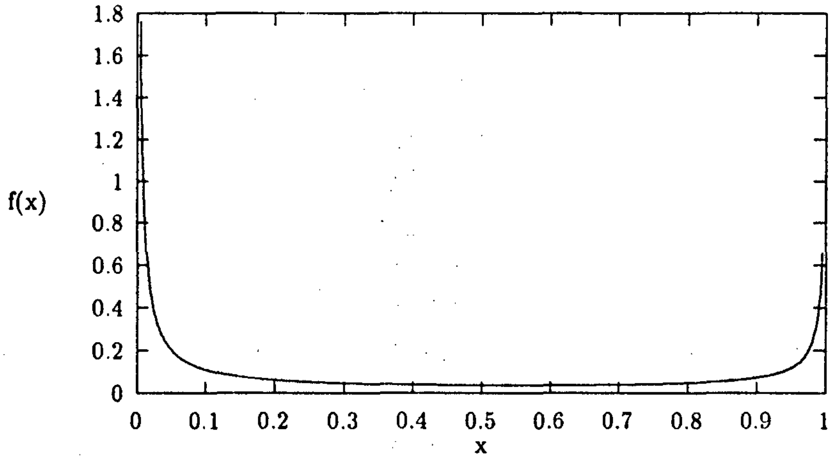

Figure 21. Distribution Function for Case 4

Is it worthwhile to attempt to expand the GLD to cover the remainder of these α_3–α_4 combinations? The answer to that question seems like it would depend on whom you asked. While mathematicians might be dismayed that we can not cover all the possible situations, simulation users might be satisfied that they can model such a wide range of different distributions using a single pdf. We tend to fall into the latter group. It is slightly disappointing that we have not been able to expand the GLD to the boundary of the Impossible Region of Figure 17, but we are happy with what we have done. The range of distributions we cannot model using the GLD seems to be "extreme" cases that will not be of much practical use. Therefore, our efforts can be considered, at least, a partial success.

VII. L-Moments for the GLD

7.1 Introduction

In determining the appropriate values for the four GLD parameters, the method of moments is commonly used. There are several reasons for this. First, the concept of the moments is something that can be understood by the majority of users. Second, the first four moments (mean, variance, skewness, and kurtosis, respectively) of a pdf or empirical data set can usually be calculated fairly easily. Third, since Ramberg and Schmeiser [12] developed equations for the first four moments as functions of the four GLD parameters, we can "match" most any distribution or data set using the GLD by simply determining appropriate values for the four λ parameters.

Unfortunately, there are some problems with this approach. As we have already mentioned, the GLD cannot "mimic" all possible combinations of skewness and kurtosis. Although it can represent most of the commonly-used pdfs, there are some it cannot. Secondly, in empirical data sets, there can be a large variability in the higher-order moments, *i.e.* the skewness and kurtosis. Since these moments are based on the third and fourth powers of the difference between each sample point and the mean, one abnormal data point can create a large change in these measures, especially when the sample data set is relatively small. It may well be the case that this outlying value is important to the overall nature of the underlying distribution, but often such occurrences are simply bad luck in sampling.

Since the GLD relies on the given values of skewness and kurtosis so heavily, the resulting distribution can be significantly altered if incorrect values are used. It therefore may be worthwhile to investigate other methods for obtaining the values of the GLD parameters. Hosking [3] has presented an alternative to the use of moments, that of L-moments. These L-moments are based on order statistics and are supposedly not as susceptible to abnormal data points as are the measures of skewness and kurtosis.

7.2 Derivation

Hosking defines the first four L-moments via:

$$\Lambda_1 = \int_0^1 R(p)dp \tag{20}$$
$$\Lambda_2 = \int_0^1 R(p)\cdot(2p-1)dp \tag{21}$$
$$\Lambda_3 = \int_0^1 R(p)\cdot(6p^2-6p+1)dp \tag{22}$$
$$\Lambda_4 = \int_0^1 R(p)\cdot(20p^3-30p^2+12p+1)dp \tag{23}$$

where R(p) is simply the percentile function.

As with the commonly-used measures of skewness and kurtosis, Hosking [3] chooses to define the two higher order moments as dimensionless ratios relative to the second order moment:

$$\tau_3 = \frac{\Lambda_3}{\Lambda_2}$$
$$\tau_4 = \frac{\Lambda_4}{\Lambda_2}.$$

We opt to follow that convention.

The four L-moments have similar roles to the typical moments. Since the first L-moment is simply the expected value of the distribution, it is identical to the mean of the distribution. Also, a symmetric distribution will have $\tau_3 = 0$, just as $\alpha_3 = 0$ for symmetric distributions using conventional moments. According to Hosking, however, τ_3 and τ_4 are more stable measures than α_3 and α_4, and therefore better estimates for empirical data.

For the GLD, the L-moment equations become:

$$\Lambda_1 = \lambda_1 + \frac{\lambda_4-\lambda_3}{\lambda_2(\lambda_3+1)(\lambda_4+1)} \tag{24}$$

$$\Lambda_2 = \frac{\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_4+1)(\lambda_4+2)} \tag{25}$$

$$\tau_3 = \frac{(\lambda_3^2-\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)-(\lambda_4^2-\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}{(\lambda_3+3)(\lambda_4+3)[\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)]} \tag{26}$$

$$\tau_4 = \frac{(\lambda_3^3-3\lambda_3^2+2\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)+(\lambda_4^3-3\lambda_4^2+2\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)}{(\lambda_3+3)(\lambda_3+4)(\lambda_4+3)(\lambda_4+4)[\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)]}. \tag{27}$$

A complete derivation of these equations can be found in appendix A.

7.3 *Usefulness of L-Moments*

Although Equations 24 through 27 look complicated, note that they are all polynomials. Unlike the GLD equations (Equations 6 through 9) based on typical moments, these have no beta functions. Since computerized solution algorithms for systems of polynomial equations are fairly common and easily adaptable, it may be easier to determine the values of the GLD parameters using L-moments instead. We have already noted that according to Hosking [3], τ_3 and τ_4 are more "stable" measures than the currently used moments, α_3 and α_4. Perhaps the method of moments could be implemented with L-moments rather than the more-familiar typical (standard) moments

7.4 *(The Problem of) Computing L-moments*

There are some problems with L-moments as well. The "sample" L-moments are based on order statistics:

$$\begin{aligned}
\Lambda_1 &= E(X)\\
\Lambda_2 &= \tfrac{1}{2}E(X_{2:2} - X_{1:2})\\
\Lambda_3 &= \tfrac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})
\end{aligned}$$

$$\Lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}),$$

where $X_{a:b}$ denotes the a^{th} order statistic in a group of size b. A review of the available literature does not show how to adapt these measures to samples of a larger size. Therefore, we have hypothesized two ways of doing this.

First, we could define each measure as the average of the results of every possible subset of the proper size within the data set. For example, in calculating the value Λ_2 for a sample of size $n = 10$, we would calculate the value of Λ_2 for all $\binom{10}{2} = 45$ possible combinations and average the results. This is obviously an unacceptable option since the amount of work required quickly becomes prohibitive, even for relatively small data sets.

A second possible method would be to derive the empirical cumulative distribution function of the data set, and create finite summations to approximate Equations 21 through 23. This seems to be the more reasonable approach, since the amount of work required is much less than that involved in calculating each possible combination. If an effective means of calculating sample L-moments can be found, we can use Equations 24 through 27 in place of the more complicated GLD moment equations to estimate appropriate values for the λ_i. Relative computational ease may simplify further research into the GLD.

Unfortunately, the quality of these estimators remains to be seen. At this point, we have no way of knowing if either method will produce "good" estimates (unbiased, minimum variance, *etc.*). To test the consistency of these approaches, it might be reasonable to conduct test cases with known pdfs to compute:

1. The theoretical L-moments—Hosking [4] presents these for many of the commonly-used pdfs.
2. An exhaustive enumeration—take samples of various sizes from the pdfs and compute.
3. The empirical cdf of those same samples.

and compare the results.

VIII. Conclusions/Suggestions for Future Research

This thesis was undertaken to study the Generalized Lambda Distribution in depth, focusing particularly on the reasons behind the limitation on the range of skewness-kurtosis combinations it can assume. Our goal was to expand this range so that the GLD would be capable of modeling any possible probability density function or empirical data set. Although we were able to expand the GLD's range beyond its previous limits, we fell short of our goal of total coverage.

What is the problem? It could be a number of things. Other general-purpose methods for fitting distributions to data (such as the Johnson and Pearson systems) must utilize more than one functional form for the distribution's pdf to cover the entire range of values shown in Figure 1. Perhaps this is the case for the GLD as well. We also could be faced with a theoretical limit on possible combinations of α_3 and α_4 that can be modeled using the GLD in its present form.

Further, since we use the method of moments to determine the appropriate values of the four GLD parameters, we limit their range to regions where the first four moments are defined. This thesis has given a description of an alternative to the method of moments (L-moments), and other methods for determining the parameter values have been documented. Perhaps one of these approaches may yield new GLD forms that the method of moments cannot.

The limit may be the result of analytical considerations, as well. Due to the complexity of the GLD, we can not solve for the parameter values in an easy fashion; instead we must rely on computerized searches to determine appropriate lambda values. The equations for α_3 and α_4, (8) and (9), both require evaluation of beta functions. The beta functions in turn require evaluation of gamma functions. Since finding exact values for the gamma function is not computationally tractable, we are instead forced to use approximation techniques. These approximations might be a source of error, but we do not believe so. However, if it is, hopefully the L-moment method, which uses only polynomials to determine the parameter values, will yield better results.

Over the course of our research, we have learned a great deal more about the GLD. We have been able, at least experimentally, to establish a lower limit on the values of kurtosis that the GLD can attain, given a specific skewness. We also have a more thorough understanding of how Powell's Algorithm works and how it functions when faced with our penalty function. We saw that neither the penalty function, nor the constrained region of viable λ_3–λ_4 values, had an effect on the the algorithm's searching process.

Where do we go from here? As a first step, the concept of L-moments is worth a longer look. At present, the literature on L–moments is limited. In particular, we do not know how to effectively determine the L-moments of an empirical data set. We discussed two of our own ideas for doing this in Chapter VII. Using L-moments, we can find the values of the four GLD parameters simply by solving a system of polynomials rather than using a non–linear function minimization required to match the standard moments. Hopefully their use can expand the GLD's range even further. L-moments are not widely known, however, and we need more information about them to determine their usefulness.

In summary, though, we must remember that even with its present restrictions, the GLD is still an extremely powerful tool for simulation studies. It allows the user to model a wide range of pdfs and empirical data sets, simply using the GLD (with appropriate parameter values) and a pseudo-random uniform random variable generator. When modeling an empirical data set, instead of facing a tough decision between two (or several) competing pdfs, we can usually match the first four moments of the data set *exactly* using the GLD—a much easier choice.

Appendix A. L-Moment Parameter Derivations

In this Appendix, we will present an in-depth derivation of the four L-moments, $\Lambda_1, \Lambda_2, \tau_3$, and τ_4 for the Generalized Lambda Distribution.

Hosking [3] presents the following formulas for generating the four L-moments from any distribution:

$$\Lambda_1 = \int_0^1 R(p)dp \tag{28}$$

$$\Lambda_2 = \int_0^1 R(p) \cdot (2p-1)dp \tag{29}$$

$$\Lambda_3 = \int_0^1 R(p) \cdot (6p^2 - 6p + 1)dp \tag{30}$$

$$\Lambda_4 = \int_0^1 R(p) \cdot (20P^3 - 30p^2 + 12p - 1)dp \tag{31}$$

where R(p) is the distribution's corresponding percentile function.

For Λ_1, we have:

$$\begin{aligned}\Lambda_1 &= \int_0^1 \left[\lambda_1 + \tfrac{1}{\lambda_2} \cdot (p^{\lambda_3} - (1-p)^{\lambda_4})\right] dp \\ &= \left[\lambda_1 p + \tfrac{1}{\lambda_2}\left(\tfrac{p^{\lambda_3+1}}{\lambda_3+1} + \tfrac{(1-p)^{\lambda_4+1}}{\lambda_4+1}\right)\right]_{p=0}^{p=1} \\ &= \lambda_1 + \tfrac{1}{\lambda_2}\left[\tfrac{1}{\lambda_3+1} - \tfrac{1}{\lambda_4+1}\right]\end{aligned}$$

which is equivalent to:

$$\Lambda_1 = \lambda_1 + \frac{\lambda_4 - \lambda_3}{\lambda_2(\lambda_3+1)(\lambda_4+1)} \tag{32}$$

For Λ_2:

$$\begin{aligned}\Lambda_2 &= \int_0^1 \left[\lambda_1 + \tfrac{1}{\lambda_2} \cdot (p^{\lambda_3} - (1-p)^{\lambda_4})\right](2p-1)dp \\ &= \int_0^1 \left[2\lambda_1 p - \lambda_1 + \tfrac{1}{\lambda_2}(2p^{\lambda_3+1} - p^{\lambda_3} - 2p(1-p)^{\lambda_4} + (1-p)^{\lambda_4})\right] dp\end{aligned}$$

$$= \left[\frac{2\lambda_1 p^2}{2} - \lambda_1 p + \frac{1}{\lambda_2}\left(\frac{2p^{\lambda_3+2}}{\lambda_3+2} - \frac{p^{\lambda_3+1}}{\lambda_3+1} - \frac{(1-p)^{\lambda_4+1}}{\lambda_4+1}\right)\right]_{p=0}^{p=1} - \frac{2}{\lambda_2}\int_0^1 p(1-p)^{\lambda_4} dp.$$

We now make use of the fact:

$$\int_0^1 x^n (1-x)^m dx = \beta(n+1, m+1) \tag{33}$$

to get:

$$\Lambda_2 = \lambda_1 - \lambda_1 + \frac{1}{\lambda_2}\left[\frac{2}{\lambda_3+2} - \frac{1}{\lambda_3+1} + \frac{1}{\lambda_4+1} - 2\beta(2, \lambda_4+1)\right].$$

But,

$$\beta(2,a) = \beta(a,2) = \frac{\Gamma(2)\Gamma(a)}{\Gamma(a+2)} = \frac{1}{a(a+1)}, \tag{34}$$

so

$$\Lambda_2 = \frac{1}{\lambda_2}\left[\frac{2}{\lambda_3+2} - \frac{1}{\lambda_3+1} + \frac{1}{\lambda_4+1} - \frac{2}{(\lambda_4+1)(\lambda_4+2)}\right].$$

This is equivalent to:

$$\Lambda_2 = \frac{2(\lambda_3+1)(\lambda_4+1)(\lambda_4+2) - (\lambda_3+2)(\lambda_4+1)(\lambda_4+2) + (\lambda_3+1)(\lambda_3+2)(\lambda_4+2) - 2(\lambda_3+1)(\lambda_3+2)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_4+1)(\lambda_4+2)} \tag{35}$$

$$\Lambda_2 = \frac{\lambda_3(\lambda_4+1)(\lambda_4+2) + \lambda_4(\lambda_3+1)(\lambda_3+2)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_4+1)(\lambda_4+2)} \tag{36}$$

For Λ_3:

$$\begin{aligned}
\Lambda_3 &= \int_0^1 \left[\lambda_1 + \frac{1}{\lambda_2}\cdot(p^{\lambda_3} - (1-p)^{\lambda_4})\right](6p^2 - 6p + 1)dp \\
&= \int_0^1 \left[6\lambda_1 p^2 - 6\lambda_1 p + \lambda_1 + \frac{1}{\lambda_2}(6p^{\lambda_3+2} - 6p^{\lambda_3+1} + p^{\lambda_3} - 6p^2(1-p)^{\lambda_4} + 6p(1-p)^{\lambda_4} - (1-p)^{\lambda_4})\right] dp \\
&= \left[\frac{6\lambda_1 p^3}{3} - \frac{6\lambda_1 p^2}{2} + \lambda_1 p\right]_{p=0}^{p=1}
\end{aligned}$$

$$+\left[\frac{1}{\lambda_2}\left(\frac{6p^{\lambda_3+3}}{\lambda_3+3}-\frac{6p^{\lambda_3+2}}{\lambda_3+2}+\frac{p^{\lambda_3+1}}{\lambda_3+1}+\frac{(1-p)^{\lambda_4+1}}{\lambda_4+1}\right)\right]_{p=0}^{p=1}-6\beta(3,\lambda_4+1)+6\beta(2,\lambda_4+1)$$

Using Equation 34 and the fact that:

$$\beta(3,a)=\beta(a,3)=\frac{\Gamma(3)\Gamma(a)}{\Gamma(a+3)}=\frac{2}{a(a+1)(a+2)}, \tag{37}$$

this equation is equivalent to:

$$\Lambda_3 = 2\lambda_1-3\lambda_1+\lambda_1$$
$$+\frac{1}{\lambda_2}\left[\frac{6}{\lambda_3+3}-\frac{6}{\lambda_3+2}+\frac{1}{\lambda_3+1}-\frac{12}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)}+\frac{6}{(\lambda_4+1)(\lambda_4+2)}-\frac{1}{\lambda_4+1}\right]$$

$$\Lambda_3 = \frac{1}{\lambda_2}\left[\frac{6}{\lambda_3+3}-\frac{6}{\lambda_3+2}+\frac{1}{\lambda_3+1}-\frac{6}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)}+\frac{6}{(\lambda_4+1)(\lambda_4+2)}-\frac{1}{\lambda_4+1}\right] \tag{38}$$

Combining the first three terms of Equation 38 yields:

$$\frac{6\lambda_3^2+18\lambda_3+12-6\lambda_3^2-24\lambda_3-18+\lambda_3^2+5\lambda_3+6}{(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}=\frac{\lambda_3^2-\lambda_3}{(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)} \tag{39}$$

Combining the last three terms of Equation 38 yields:

$$\frac{-12+6(\lambda_4+3)-(\lambda_4+2)(\lambda_4+3)}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)}=\frac{-\lambda_4^2+\lambda_4}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)} \tag{40}$$

We then combine the results of Equations 39 and 40 to get:

$$\Lambda_3 = \frac{1}{\lambda_2}\left[\frac{\lambda_3^2-\lambda_3}{(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}-\frac{\lambda_4^2-\lambda_4}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)}\right]$$

$$= \frac{(\lambda_3^2-\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)-(\lambda_4^2-\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)} \tag{41}$$

We use Equations 36 and 41 to derive the equation for τ_3:

$$\tau_3 = \frac{\Lambda_3}{\Lambda_2}$$

$$\tau_3 = \frac{(\lambda_3^2-\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)-(\lambda_4^2-\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)} \cdot \frac{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_4+1)(\lambda_4+2)}{\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)}$$

$$\tau_3 = \frac{(\lambda_3^2-\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)-(\lambda_4^2-\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)}{(\lambda_3+3)(\lambda_4+3)[\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)]} \qquad (42)$$

For Λ_4:

$$\begin{aligned}
\Lambda_4 &= \int_0^1 \left[\lambda_1 + \tfrac{1}{\lambda_2}\cdot(p^{\lambda_3} - (1-p)^{\lambda_4})\right](20p^3 - 30p^2 + 12p - 1)dp \\
&= \int_0^1 20\lambda_1 p^3 - 30\lambda_1 p^2 + 12\lambda_1 p - \lambda_1 \\
&\quad + \tfrac{1}{\lambda_2}(20p^{\lambda_3+3} - 30p^{\lambda_3+2} + 12p^{\lambda_3+1} - p^{\lambda_3} \\
&\quad -20p^3(1-p)^{\lambda_4} + 30p^2(1-p)^{\lambda_4} - 12p(1-p)^{\lambda_4}) + (1-p)^{\lambda_4})dp \\
&= \left[\tfrac{20\lambda_1 p^4}{4} - \tfrac{30\lambda_1 p^3}{3} + \tfrac{12\lambda_1 p^2}{2} - \lambda_1 p\right]_{p=0}^{p=1} \\
&\quad + \tfrac{1}{\lambda_2}\left[\tfrac{20p^{\lambda_3+4}}{\lambda_3+4} - \tfrac{30p^{\lambda_3+3}}{\lambda_3+3} + \tfrac{12p^{\lambda_3+2}}{\lambda_3+2} - \tfrac{p^{\lambda_3+1}}{\lambda_3+1} - \tfrac{(1-p)^{\lambda_4+1}}{\lambda_4+1}\right]_{p=0}^{p=1} \\
&\quad + \tfrac{1}{\lambda_2}[-20\beta(4,\lambda_4+1) + 30\beta(3,\lambda_4+1) - 12\beta(2,\lambda_4+1)]
\end{aligned}$$

We use Equations 34, 37 and the fact that:

$$\beta(4,a) = \beta(a,4) = \frac{\Gamma(4)\Gamma(a)}{\Gamma(a+4)} = \frac{6}{a(a+1)(a+2)(a+3)},$$

to find an equivalent expression.

$$\Lambda_4 = \quad 5\lambda_1 - 10\lambda_1 + 6\lambda_1 - \lambda_1 + \frac{1}{\lambda_2}\left[\frac{20}{\lambda_3+4} - \frac{30}{\lambda_3+3} + \frac{12}{\lambda_3+2} - \frac{1}{\lambda_3+1}\right]$$

$$+\frac{1}{\lambda_2}\left[-\frac{120}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)} + \frac{60}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)} - \frac{12}{(\lambda_4+1)(\lambda_4+2)} + \frac{1}{\lambda_4+1}\right]$$

$$\Lambda_4 = \quad \frac{1}{\lambda_2}\left[\frac{20}{\lambda_3+4} - \frac{30}{\lambda_3+3} + \frac{12}{\lambda_3+2} - \frac{1}{\lambda_3+1}\right]$$

$$+\frac{1}{\lambda_2}\left[-\frac{120}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)} + \frac{60}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)} - \frac{12}{(\lambda_4+1)(\lambda_4+2)} + \frac{1}{\lambda_4+1}\right] \qquad (43)$$

Combining the first four terms of Equation 43 yields:

$$\frac{1}{D}\left[20(\lambda_3+1)(\lambda_3+2)(\lambda_3+3) - 30(\lambda_3+1)(\lambda_3+2)(\lambda_3+4)\right]$$

$$+ \quad \frac{1}{D}\left[12(\lambda_3+1)(\lambda_3+3)(\lambda_3+4) - (\lambda_3+2)(\lambda_3+3)(\lambda_3+4)\right]$$

where D represents the common denominator:

$$D = (\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4).$$

This is equivalent to:

$$\frac{\lambda_3^3 - 3\lambda_3^2 + 2\lambda_3}{(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)}. \qquad (44)$$

Combining the last four terms of Equation 43 yields:

$$\frac{-120+60(\lambda_4+4)-12(\lambda_4+3)(\lambda_4+4)+(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)}$$

$$= \quad \frac{\lambda_4^3-3\lambda_4^2+2\lambda_4}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)}. \qquad (45)$$

Taking Equations 44 and 45 together yields:

$$\Lambda_4 = \frac{1}{\lambda_2}\left[\frac{\lambda_3^3-3\lambda_3^2+2\lambda_3}{(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)} + \frac{\lambda_4^3-3\lambda_4^2+2\lambda_4}{(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)}\right]$$

$$\Lambda_4 = \frac{(\lambda_3^3-3\lambda_3^2+2\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)+(\lambda_4^3-3\lambda_4^2+2\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)}. \tag{46}$$

Using Equations 36 and 46, we find:

$$\begin{aligned}\tau_4 &= \frac{\Lambda_4}{\Lambda_2} \\ &= \frac{(\lambda_3^3-3\lambda_3^2+2\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)+(\lambda_4^3-3\lambda_4^2+2\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)} \\ &\quad \cdot \frac{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_4+1)(\lambda_4+2)}{\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)}\end{aligned}$$

$$\tau_4 = \frac{(\lambda_3^3-3\lambda_3^2+2\lambda_3)(\lambda_4+1)(\lambda_4+2)(\lambda_4+3)(\lambda_4+4)+(\lambda_4^3-3\lambda_4^2+2\lambda_4)(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)}{(\lambda_3+3)(\lambda_3+4)(\lambda_4+3)(\lambda_4+4)[\lambda_3(\lambda_4+1)(\lambda_4+2)+\lambda_4(\lambda_3+1)(\lambda_3+2)]}. \tag{47}$$

Bibliography

1. Karl V. Bury. *Statistical Models in Applied Science.* John Wiley and Sons, 1975.
2. J. J. Filliben. *Simple and Robust Linear Estimation of the Location Parameters of a Symmetric Distribution.* PhD thesis, Princeton University, 1969.
3. J. R. M. Hosking. L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society*, Ser B, 52:105–124, 1990.
4. J. R. M. Hosking. Moments or L-moments? An Example Comparing Two Measures of Distributional Shape. *The American Statistician*, 46(3):186–189, August 1992.
5. Chung-Lung Hsu. A User Friendly Software Package for Determining the Parameters of the Generalized Lambda Distribution. Master's thesis, Auburn University, 1991.
6. Zaven A. Karian and Edward J. Dudewicz. *Modern Statisical, Systems, and GPSS Simulation: The First Course.* Computer Science Press, 1991.
7. James L. Kuester and Joe H. Mize. *Optimization Techniques With FORTRAN.* McGraw-Hill, Inc., 1973.
8. William Mendenhall, Dennis D. Wackerly, and Richard L. Scheaffer. *Mathematical Statistics with Applications.* PWS-Kent Publishing, 1990.
9. Edward F. Mykytka. Some Useful Properties and Methods for Determining the Parameters of the Ramberg-Schmeiser-Tukey Distribution. Master's thesis, University of Iowa, 1978.
10. Edward F. Mykytka and John S. Ramberg. Method of Moments Estimators For the Generalized Lambda Distribution. Article Awaiting Publication.
11. M. J. D. Powell. An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives. *Computer Journal*, 7:155–162, 1964.
12. John S. Ramberg and Bruce W. Schmeiser. An Approximate Method for Generating Symmetric Random Variables. *Communications of the ACM*, 15:987–990, 1972.
13. John S. Ramberg, Pandu R. Tadikamalla, Edward J. Dudewicz, and Edward F. Mykytka. A Probability Distribution and Its Uses in Fitting Data. *Technometrics*, 21(2):201–214, 1979.
14. Bruce W. Schmeiser. A General Algorithm for Generating Random Variables. Master's thesis, University of Iowa, 1971.
15. J. W. Tukey. The Practical Relationship Between Common Transformations of Percentages of Counts and of Amounts. Technical report, Princeton Univeristy Statistical Techniques Research Group, 1960.

Vita

First Lieutenant Robert J. Bergevin was born on 14 April 1967 in Plattsburgh, NY. He attended Plattsburgh High School, where he was named salutatorian of his graduating class. He earned the degree of Bachelor of Science in Electrical Engineering from the Massachusetts Institute of Technology (MIT) in June 1989. He received his commission through MIT's Air Force Reserve Officer Training Corps and briefly worked as a Systems Engineer in the Communication System Segment Replacement (CSSR) Program Office at Hanscom AFB, MA before attending Undergraduate Pilot Training (UPT) at Williams AFB, AZ in May 1990. In May 1991, he graduated from UPT and was assigned to the Air Force Institute of Technology, where he earned the degree of Master of Science in Operations Research. In March 1993, he was assigned to the Rated Force Analysis Office of the Air Force Military Personnel Center (AFMPC) at Randolph AFB, TX, while awaiting a flying assignment.

Permanent address: 28 Prospect Ave
Plattsburgh, NY 12901

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	MARCH 1993	Master's Thesis

4. TITLE AND SUBTITLE
An Analysis of the Generalized Lambda Distribution

5. FUNDING NUMBERS

6. AUTHOR(S)
1LT Robert J. Bergevin, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GST/ENS/93M-01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
~~ENS/AFIT~~
~~Wright-Patterson AFB, OH 45433~~

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
UNLIMITED DISTRIBUTION

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)* The Generalized Lambda Distribution (GLD) is a four parameter function that is capable of mimicking the behavior of a wide range of probability density functions (pdfs). Unfortunately, the GLD presently cannot model every possible type of pdf. Since the reasons for this limitation are unknown, this thesis examines several potential problems in an attempt to expand the range of distributions the GLD can mimic.

We first present a discussion of the behavior of the algorithm that is used to search for the appropriate GLD parameter values. In particular, we examine the effect of using an unconstrained search to find the parameters subject to a constraint that ensures that the resulting pdf is valid. We also develop a reparameterization of the GLD that creates an unconstrained search region. This does not expand the range of distributions the GLD can mimic.

We then use an extensive numerical investigation to examine the set of distributions that can be obtained from combinations of the GLD parameters. This examination allows us to expand the range of pdfs that the GLD can model. We also inspect some pdfs that cannot be modeled using the GLD, as well as present an alternative to the method of moments for determining parameter values, using the concept of L-moments.

14. SUBJECT TERMS
Generalized Lambda Distribution (GLD), Powell's Algorithm, Method of Moments, L-moments

15. NUMBER OF PAGES
80

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

END

FILMED

DATE: 4-93

DTIC